

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2004, 21:00:56 ; Search time 50 Seconds
(without alignments)
885.694 Million cell updates/sec

Title: US-10-014-927-19

Perfect score: 1451

Sequence: 1 MSSRWRTIYVGNLPGDIRK.....RSKSRSRSRSPVSPVWISG 279

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 159726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1377.5	94.9	268	21	AA11416
2	1359.5	93.7	264	21	AA11416
3	1359.5	93.7	270	21	AA11416
4	1343	92.6	256	21	AA11416
5	1343	92.6	256	21	AA11416
6	1159	79.9	237	21	AA11416
7	865	59.6	303	21	AA11416
8	863	59.5	307	21	AA11416
9	836	57.6	303	21	AA11416

10	805	55.5	296	21	AA115601	Arabidopsis thalia
11	663.5	45.7	136	21	AA115601	Arabidopsis thalia
12	663.5	45.7	142	21	AA115601	Arabidopsis thalia
13	647	44.6	128	21	AA115601	Arabidopsis thalia
14	646.5	44.6	294	21	AA115601	Arabidopsis thalia
15	582.5	40.1	248	21	AA115601	Human SF2/ASF prot
16	582.5	40.1	248	21	AA115601	Human shear stress
17	547	37.7	255	22	AA115601	Drosophila melanog
18	517	35.6	253	21	AA115601	Human prostate can
19	517	35.6	254	23	AA115601	Human ovarian anti
20	472	32.5	124	21	AA115601	Zea mays protein f
21	472	32.5	159	21	AA115601	Zea mays protein f
22	471.5	32.5	125	21	AA115601	Zea mays protein f
23	471.5	32.5	154	21	AA115601	Zea mays protein f
24	457.5	31.5	99	21	AA115601	Arabidopsis thalia
25	453	31.2	227	22	AA115601	Novel human secret
26	436	30.0	96	21	AA115601	Arabidopsis thalia
27	431	29.7	329	22	AA115601	Drosophila melanog
28	411.5	28.4	583	22	AA115601	Novel human secret
29	391	26.9	83	21	AA115601	Zea mays protein f
30	374.5	25.8	156	21	AA115601	Arabidopsis thalia
31	363	25.0	165	21	AA115601	Arabidopsis thalia
32	363	25.0	165	21	AA115601	Arabidopsis thalia
33	356.5	24.6	155	21	AA115601	Arabidopsis thalia
34	352	24.3	379	22	AA115601	Novel human secret
35	339	23.4	153	22	AA115601	Novel human secret
36	309	21.3	278	22	AA115601	Novel human diagno
37	306	21.1	297	22	AA115601	Novel human secret
38	277	19.1	306	21	AA115601	Human pancreatic c
39	275.5	19.0	94	23	AA115601	Human ORF172 prote
40	269.5	18.6	96	23	AA115601	Human ORFX protein
41	268.5	18.5	117	21	AA115601	Arabidopsis thalia
42	258.5	17.8	284	21	AA115601	Arabidopsis thalia
43	258.5	17.8	313	21	AA115601	Arabidopsis thalia
44	256	17.6	290	21	AA115601	Arabidopsis thalia
45	254.5	17.5	283	22	AA115601	Human polypeptide

ALIGNMENTS

RESULT 1	
AA11416	
ID	AA11416 standard; Protein; 268 AA.
AC	AA11416;
DT	23-FEB-2001 (first entry)
DE	A. thaliana SRP30 protein.
KW	SR protein; splice-factor activity; plant; developmental behavior;
KW	flowering; crop plant; cereal; rice; fruit.
OS	Arabidopsis thaliana.
PN	WO200065059-A1.
PD	02-NOV-2000.
PF	20-APR-2000; 2000WO-AT00100.
PR	23-APR-1999; 99AT-0000727.
PA	(OSTP) OESTERR FORSCH SEIBERSDORF.
PI	Barta A, Lopato S, Kalyana M, Dorner S;
DR	WPI; 2000-687349/67.
PT	Novel proteins with splice-factor activity in plants, useful e.g. for altering flowering time or development, and the nucleic acid that encodes it -

```

XX Disclosure; Fig 2; 67pp; German.
PS This invention describes a novel protein (I) with splice-factor activity
CC in plants (I) modifies the choice of splice sites in many plant
CC pre-mRNAs. (I) (also the nucleic acid that encodes them and related
CC vectors or expression systems) are used: (i) to alter splice patterns in
CC plants, or their parts; (ii) to alter developmental behavior of plants;
CC and/or (iii) to delay flowering, particularly by at least 25% relative
CC to the wild type, especially in crop plants such as cereals, beans, rice
CC and fruit.
XX
SQ Sequence 268 AA;
    Query Match          94.9%; Score 1377.5; DB 21; Length 268;
    Best Local Similarity 96.1%; Pred. No. 8.5e-128;
    Matches 268; Conservative 0; Mismatches 0; Indels 11; Gaps 1;
QY 1 MSSRWNTIYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPPPPGYAFVEFEDPRDAD 60
DB 1 MSSRWNTIYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPPPPGYAFVEFEDPRDAD 60
QY 61 AIYGRDGYDFGCLRLVEIAHGRRFSPVDRYSSSYASGRAPSRSDYRVLVTGLPPSA 120
DB 61 AIYGRDGYDFGCLRLVEIAHGRRFSPVDRYSSSYASGRAPSRSDYRVLVTGLPPSA 120
QY 121 SWQDKHMRKAGVCFSEVPPDRKMGSGVVDYSDYDMKYAIRKLDATFERNAPSSAYI 180
DB 121 SWQDKHMRKAGVCFSEVPPDRKMGSGVVDYSDYDMKYAIRKLDATFERNAPSSAYI 180
QY 181 RVRYEGRSVSRPDDSKSVESRSGRSPGSCYSKSRSVSPARSISPRRPLSRGRSLY 240
DB 181 RVRYEGRSVSRPDDSKSVESRSGRSPGSCYSKSRSVSPARSISPRRPLSRGRSLY 240
QY 241 SSVSRSGSLLAGDWISQSRKSRSRSRNSRNSPVSFVIG 279
DB 241 SSVSR-----SQSRKSRSRSRNSRNSPVSFVIG 268

RESULT 2
AAG47934
ID AAG47934 standard; Protein; 264 AA.
AC AAG47934;
XX
XX
XX
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 60472.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123549.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127463.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
XX 28-JUN-1999; 99US-0140823.
XX 29-JUN-1999; 99US-0140991.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
XX 01-JUL-1999; 99US-0142154.
XX 02-JUL-1999; 99US-0142055.
XX 06-JUL-1999; 99US-0142390.
XX 08-JUL-1999; 99US-0142803.
XX 09-JUL-1999; 99US-0142920.
XX 12-JUL-1999; 99US-0142977.
XX 13-JUL-1999; 99US-0143542.
XX 14-JUL-1999; 99US-0143624.
XX 15-JUL-1999; 99US-0144005.
XX 16-JUL-1999; 99US-0144085.
XX 16-JUL-1999; 99US-0144086.
XX 19-JUL-1999; 99US-0144325.
XX 19-JUL-1999; 99US-0144331.
XX 19-JUL-1999; 99US-0144332.
XX 19-JUL-1999; 99US-0144333.
XX 19-JUL-1999; 99US-0144334.
XX 19-JUL-1999; 99US-0144335.
XX 20-JUL-1999; 99US-0144352.

```



```
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159299.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.

PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 93.7%; Score 1359.5; DB 21; Length 264;
Best Local Similarity 96.0%; Pred. No. 5e-126;
Matches 264; Conservative 0; Mismatches 0; Indels 11; Gaps 1;

QY 1 MMSRNRRTIYVGNLPGDIRKCEVEDLFYKYGPIVVIDLKIPRPPGYAFVEFEDPRDADD 60
Db 1 MMSRNRRTIYVGNLPGDIRKCEVEDLFYKYGPIVVIDLKIPRPPGYAFVEFEDPRDADD 60
QY 61 AIYGRDGYDFDCCRLRVEIAHGRRFSPSVDRYSSYSASRAPSRSDYEVLTGLPPSA 120
Db 61 AIYGRDGYDFDCCRLRVEIAHGRRFSPSVDRYSSYSASRAPSRSDYEVLTGLPPSA 120
QY 121 SWQDLKDHMRKAGDVCFSEVFPDRKMGVGVVDYSDYKAIKLDATFENAFSSAYI 180
Db 121 SWQDLKDHMRKAGDVCFSEVFPDRKMGVGVVDYSDYKAIKLDATFENAFSSAYI 180
QY 181 RVREYESRSVSRPDDSKSYRSRSGPSCSYSSKRSVSPARSISPRSLRSRSY 240
Db 181 RVREYESRSVSRPDDSKSYRSRSGPSCSYSSKRSVSPARSISPRSLRSRSY 240
QY 241 SSVSRGSLLRAGDWISQSRKSRGRSRNSPVSP 275
Db 241 SSVSR-----SQSRKSRGRSRNSPVSP 264

RESULT 3
AAG23883
ID AAG23883 standard; Protein; 270 AA.
XX
AC AAG23883,
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27353.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW Hybridisation assay; Genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
```

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139482.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142997.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

```
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 93.7%; Score 1359.5; DB 21; Length 270;
Best Local Similarity 95.0%; Pred. No. 5.2e-126;
Matches 267; Conservative 0; Mismatches 1; Indels 13; Gaps 2;

QY 1 MSSRWNTIYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPRPPGYAFVEFEDPRDADD 60
Db 1 MSSRWNTIYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPRPPGYAFVEFEDPRDADD 60
QY 61 AIYGRDGYDFDGCRLRVEIAHGRRRPSVDYSSYSASRAPSRSDYRVLTGLPPSA 120
Db 61 AIYGRDGYDFDGCRLRVEIAHGRRRPSVDYSSYSASRAPSRSDYRVLTGLPPSA 120
QY 121 SWQDLKDHMRKAGDVCFSEVFPDRKMGSGVVDYSDYDMKYAIRKLDATFENAFSSAYI 180
Db 121 SWQDLKDHMRKAGDVCFSEVFPDRKMGSGVVDYSDYDMKYAIRKLDATFENAFSSAYI 180
QY 181 RVREYESRSVSRSPDDSKYSRSRSGPSCSYSSKSRSVSPARSTSPSRPLSRSRSLY 240
Db 181 RVREYESRSVSRSPDDSKYSRSRSGPSCSYSSKSRSVSPARSTSPSRPLSRSRSPY 240
QY 241 SSVSRSGSLLRAGDWI 256
Db 241 SSVSRSGSLLRAGDWI 256

RESULT 4
AAB11414
ID AAB11414 standard; Protein; 256 AA.
XX
XX AAB11414;
AC
XX
XX 23-FEB-2001 (first entry)
DT
XX
XX A. thaliana SRp30 protein.
DE
XX
XX SR protein; splice-factor activity; plant; developmental behavior;
KW
XX flowering; crop plant; cereal; bean; rice; fruit.
XX
XX Arabidopsis thaliana.
OS
XX
XX WO200065059-A1.
PN
```

```
XX 02-NOV-2000.
XX
XX 20-APR-2000; 2000WO-AT00100.
XX
XX 23-APR-1999; 99AT-0000727.
XX
XX (OSTP ) OESTERR FORSCH SEIBERSDORF.
XX
XX Barta A, Lopato S, Kalyna M, Dorner S;
XX WPI; 2000-687349/67.
XX N-PSDB; AAC81899.
XX
XX Novel proteins with splice-factor activity in plants, useful e.g. for
XX altering flowering time or development, and the nucleic acid that
XX encodes it -
XX
XX Claim 1; Fig 1A; 67pp; German.
XX
XX This invention describes a novel protein (I) with splice-factor activity
XX in plants (I) modifies the choice of splice sites in many plant
XX pre-mRNAs. (I) (also the nucleic acid that encodes them and related
XX vectors or expression systems) are used: (i) to alter splice patterns in
XX plants, or their parts; (ii) to alter developmental behavior of plants;
XX and/or (iii) to delay flowering, particularly by at least 25% relative
XX to the wild type, especially in crop plants such as cereals, beans, rice
XX and fruit.
XX
XX Sequence 256 AA;

Query Match 92.6%; Score 1343; DB 21; Length 256;
Best Local Similarity 100.0%; Pred. No. 2.1e-124;
Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSRWNTIYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPRPPGYAFVEFEDPRDADD 60
Db 1 MSSRWNTIYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPRPPGYAFVEFEDPRDADD 60
QY 61 AIYGRDGYDFDGCRLRVEIAHGRRRPSVDYSSYSASRAPSRSDYRVLTGLPPSA 120
Db 61 AIYGRDGYDFDGCRLRVEIAHGRRRPSVDYSSYSASRAPSRSDYRVLTGLPPSA 120
QY 121 SWQDLKDHMRKAGDVCFSEVFPDRKMGSGVVDYSDYDMKYAIRKLDATFENAFSSAYI 180
Db 121 SWQDLKDHMRKAGDVCFSEVFPDRKMGSGVVDYSDYDMKYAIRKLDATFENAFSSAYI 180
QY 181 RVREYESRSVSRSPDDSKYSRSRSGPSCSYSSKSRSVSPARSTSPSRPLSRSRSLY 240
Db 181 RVREYESRSVSRSPDDSKYSRSRSGPSCSYSSKSRSVSPARSTSPSRPLSRSRSLY 240
QY 241 SSVSRSGSLLRAGDWI 256
Db 241 SSVSRSGSLLRAGDWI 256

RESULT 5
AAG47918
ID AAG47918 standard; Protein; 256 AA.
XX
XX AAG47918;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 60450.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
```

[illegible]

PR	02-JUL-1999;	99US-01420559;
PR	06-JUL-1999;	99US-01423093;
PR	08-JUL-1999;	99US-01428033;
PR	09-JUL-1999;	99US-01429200;
PR	12-JUL-1999;	99US-01429777;
PR	13-JUL-1999;	99US-0143542;
PR	14-JUL-1999;	99US-01436241;
PR	15-JUL-1999;	99US-01440025;
PR	16-JUL-1999;	99US-01440855;
PR	16-JUL-1999;	99US-01440886;
PR	19-JUL-1999;	99US-01443255;
PR	19-JUL-1999;	99US-01443311;
PR	19-JUL-1999;	99US-01443322;
PR	19-JUL-1999;	99US-01443333;
PR	19-JUL-1999;	99US-01443334;
PR	19-JUL-1999;	99US-01443335;
PR	20-JUL-1999;	99US-01443352;
PR	20-JUL-1999;	99US-0144632;
PR	20-JUL-1999;	99US-0144881;
PR	21-JUL-1999;	99US-01448411;
PR	21-JUL-1999;	99US-01450886;
PR	21-JUL-1999;	99US-01450887;
PR	22-JUL-1999;	99US-01450885;
PR	22-JUL-1999;	99US-01450877;
PR	22-JUL-1999;	99US-01450878;
PR	22-JUL-1999;	99US-01450879;
PR	22-JUL-1999;	99US-01451922;
PR	23-JUL-1999;	99US-01451945;
PR	23-JUL-1999;	99US-01452181;
PR	23-JUL-1999;	99US-01452284;
PR	26-JUL-1999;	99US-01452766;
PR	27-JUL-1999;	99US-01459133;
PR	27-JUL-1999;	99US-01459183;
PR	28-JUL-1999;	99US-01459199;
PR	28-JUL-1999;	99US-01459519;
PR	02-AUG-1999;	99US-01463865;
PR	02-AUG-1999;	99US-01463888;
PR	02-AUG-1999;	99US-01463899;
PR	03-AUG-1999;	99US-01470338;
PR	04-AUG-1999;	99US-01472004;
PR	04-AUG-1999;	99US-01473002;
PR	05-AUG-1999;	99US-01471922;
PR	05-AUG-1999;	99US-01472600;
PR	06-AUG-1999;	99US-01473003;
PR	06-AUG-1999;	99US-01474163;
PR	09-AUG-1999;	99US-01474933;
PR	09-AUG-1999;	99US-01479335;
PR	10-AUG-1999;	99US-01481771;
PR	11-AUG-1999;	99US-01483193;
PR	12-AUG-1999;	99US-01483411;
PR	13-AUG-1999;	99US-01485655;
PR	13-AUG-1999;	99US-01486884;
PR	16-AUG-1999;	99US-01493688;
PR	17-AUG-1999;	99US-01491755;
PR	18-AUG-1999;	99US-01494266;
PR	20-AUG-1999;	99US-01497222;
PR	20-AUG-1999;	99US-01497323;
PR	20-AUG-1999;	99US-01499299;
PR	23-AUG-1999;	99US-01499902;
PR	23-AUG-1999;	99US-01499300;
PR	26-AUG-1999;	99US-01505666;
PR	27-AUG-1999;	99US-01508884;
PR	27-AUG-1999;	99US-01510655;
PR	27-AUG-1999;	99US-01510666;
PR	27-AUG-1999;	99US-01510800;
PR	30-AUG-1999;	99US-01513003;
PR	31-AUG-1999;	99US-01514338;
PR	07-SEP-1999;	99US-01523263;
PR	10-SEP-1999;	99US-01530758;
PR	13-SEP-1999;	99US-01537588;
PR	15-SEP-1999;	99US-01540188;
PR	16-SEP-1999;	99US-01540399;
PR	20-SEP-1999;	99US-01547799;

DE	Arabidopsis thaliana protein fragment SEQ ID NO: 37145.
XX	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KM	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
XX	25-FEB-2000; 2000EP-0301439.
PF	
XX	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0128845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	23-APR-1999; 99US-0130891.
PR	28-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0132407.
PR	04-MAY-1999; 99US-0132484.
PR	05-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	06-MAY-1999; 99US-0132487.
PR	07-MAY-1999; 99US-0132863.
PR	11-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134221.
PR	14-MAY-1999; 99US-0134370.
PR	18-MAY-1999; 99US-0134768.
PR	19-MAY-1999; 99US-0134941.
PR	20-MAY-1999; 99US-0135124.
PR	21-MAY-1999; 99US-0135353.
PR	24-MAY-1999; 99US-0135629.
PR	25-MAY-1999; 99US-0136021.
PR	27-MAY-1999; 99US-0136392.
PR	28-MAY-1999; 99US-0136782.
PR	01-JUN-1999; 99US-0137222.
PR	03-JUN-1999; 99US-0137528.
PR	04-JUN-1999; 99US-0137502.
PR	07-JUN-1999; 99US-0137724.
PR	08-JUN-1999; 99US-0138094.
PR	10-JUN-1999; 99US-0138540.
PR	10-JUN-1999; 99US-0138847.
PR	14-JUN-1999; 99US-0139119.
PR	16-JUN-1999; 99US-0139452.
PR	16-JUN-1999; 99US-0139453.
PR	17-JUN-1999; 99US-0139492.
PR	18-JUN-1999; 99US-0139454.
PR	18-JUN-1999; 99US-0139455.
PR	18-JUN-1999; 99US-0139456.
PR	18-JUN-1999; 99US-0139457.
PR	18-JUN-1999; 99US-0139458.
PR	18-JUN-1999; 99US-0139459.
PR	18-JUN-1999; 99US-0139460.
PR	18-JUN-1999; 99US-0139461.
PR	18-JUN-1999; 99US-0139462.
PR	18-JUN-1999; 99US-0139463.
PR	18-JUN-1999; 99US-0139750.
PR	18-JUN-1999; 99US-0139763.
PR	21-JUN-1999; 99US-0139817.

	PR	22-SEP-1999;	99US-01551139.	
	PR	23-SEP-1999;	99US-0155486.	
	PR	24-SEP-1999;	99US-0155659.	
	PR	28-SEP-1999;	99US-0156458.	
	PR	29-SEP-1999;	99US-0156596.	
	PR	04-OCT-1999;	99US-0157117.	
	PR	05-OCT-1999;	99US-0157753.	
	PR	06-OCT-1999;	99US-0157865.	
	PR	07-OCT-1999;	99US-0158029.	
	PR	08-OCT-1999;	99US-0158232.	
	PR	12-OCT-1999;	99US-0158369.	
	PR	13-OCT-1999;	99US-0159293.	
	PR	13-OCT-1999;	99US-0159294.	
	PR	13-OCT-1999;	99US-0159295.	
	PR	14-OCT-1999;	99US-0159329.	
	PR	14-OCT-1999;	99US-0159330.	
	PR	14-OCT-1999;	99US-0159331.	
	PR	14-OCT-1999;	99US-0159637.	
	PR	14-OCT-1999;	99US-0160770.	
	PR	18-OCT-1999;	99US-0159638.	
	PR	18-OCT-1999;	99US-0159584.	
	PR	21-OCT-1999;	99US-0160741.	
	PR	21-OCT-1999;	99US-0160767.	
	PR	21-OCT-1999;	99US-0160768.	
	PR	21-OCT-1999;	99US-0160770.	
	PR	21-OCT-1999;	99US-0160814.	
	PR	21-OCT-1999;	99US-0160815.	
	PR	22-OCT-1999;	99US-0160980.	
	PR	22-OCT-1999;	99US-0160981.	
	PR	22-OCT-1999;	99US-0160989.	
	PR	25-OCT-1999;	99US-0161404.	
	PR	25-OCT-1999;	99US-0161405.	
	PR	25-OCT-1999;	99US-0161406.	
	PR	26-OCT-1999;	99US-0161359.	
	PR	26-OCT-1999;	99US-0161360.	
	PR	26-OCT-1999;	99US-0161361.	
	PR	28-OCT-1999;	99US-0161920.	
	PR	28-OCT-1999;	99US-0161920.	
	PR	28-OCT-1999;	99US-0161993.	
	PR	29-OCT-1999;	99US-0162142.	

Query Match 92.6%; Score 1343; DB 21; Length 256;
 Best Local Similarity 100.0%; Pred. No. 2.le-124;
 Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY	1	MSSRNRTIYVGNLPGDIRKCEVEDLFYKYGPVIDIDLKIPPPGCGAFVEFEDPRDADD	60
DB	1	MSSRNRTIYVGNLPGDIRKCEVEDLFYKYGPVIDIDLKIPPPGCGAFVEFEDPRDADD	60
QY	61	AIYGRDGVDFGCRLRVEIAHGRRFPSPSDRYSSSYASAPRSRDYRVLTGLPSPA	120
DB	61	AIYGRDGVDFGCRLRVEIAHGRRFPSPSDRYSSSYASAPRSRDYRVLTGLPSPA	120
QY	121	SMQDLKHNRKAGDVCFSEVPDPDKMGSGVVDSNVDDMKYAIRKLDATEFRNAFSSAYI	180
DB	121	SMQDLKHNRKAGDVCFSEVPDPDKMGSGVVDSNVDDMKYAIRKLDATEFRNAFSSAYI	180
QY	181	RVREYESRSVSPPDCKSKYSRSRSPSCPSYSSKSRSVSPARSTSPSRPLSRSRSLY	240
DB	181	RVREYESRSVSPPDCKSKYSRSRSPSCPSYSSKSRSVSPARSTSPSRPLSRSRSLY	240
QY	241	SSVSRSGSLLRAGDWI	256
DB	241	SSVSRSGSLLRAGDWI	256

RESULT 6
 AAG30990
 ID AAG30990 standard; Protein; 237 AA.
 XX
 AC AAG30990;
 XX
 DT 17-OCT-2000 (first entry)
 YY

PR 22-JUN-1999;	99US-0139899;	PR 30-AUG-1999;	99US-0151303;
PR 23-JUN-1999;	99US-0140353;	PR 31-AUG-1999;	99US-0151438;
PR 23-JUN-1999;	99US-0140354;	PR 01-SEP-1999;	99US-0151930;
PR 24-JUN-1999;	99US-0140695;	PR 07-SEP-1999;	99US-0152363;
PR 28-JUN-1999;	99US-0140823;	PR 10-SEP-1999;	99US-0153070;
PR 29-JUN-1999;	99US-0140991;	PR 13-SEP-1999;	99US-0153758;
PR 30-JUN-1999;	99US-0141287;	PR 15-SEP-1999;	99US-0154018;
PR 01-JUL-1999;	99US-0141842;	PR 20-SEP-1999;	99US-0154039;
PR 01-JUL-1999;	99US-0142154;	PR 22-SEP-1999;	99US-0154779;
PR 02-JUL-1999;	99US-0142055;	PR 23-SEP-1999;	99US-0155139;
PR 06-JUL-1999;	99US-0142390;	PR 23-SEP-1999;	99US-0155486;
PR 08-JUL-1999;	99US-0142803;	PR 24-SEP-1999;	99US-0156659;
PR 09-JUL-1999;	99US-0142920;	PR 28-SEP-1999;	99US-0156458;
PR 12-JUL-1999;	99US-0142977;	PR 29-SEP-1999;	99US-0156596;
PR 13-JUL-1999;	99US-0143542;	PR 04-OCT-1999;	99US-0157117;
PR 14-JUL-1999;	99US-0143624;	PR 05-OCT-1999;	99US-0157753;
PR 15-JUL-1999;	99US-0144005;	PR 06-OCT-1999;	99US-0157865;
PR 16-JUL-1999;	99US-0144085;	PR 07-OCT-1999;	99US-0158023;
PR 16-JUL-1999;	99US-0144086;	PR 08-OCT-1999;	99US-0158232;
PR 19-JUL-1999;	99US-0144325;	PR 12-OCT-1999;	99US-0158369;
PR 19-JUL-1999;	99US-0144331;	PR 13-OCT-1999;	99US-0159293;
PR 19-JUL-1999;	99US-0144332;	PR 13-OCT-1999;	99US-0159294;
PR 19-JUL-1999;	99US-0144333;	PR 13-OCT-1999;	99US-0159295;
PR 19-JUL-1999;	99US-0144334;	PR 14-OCT-1999;	99US-0159323;
PR 19-JUL-1999;	99US-0144335;	PR 14-OCT-1999;	99US-0159330;
PR 20-JUL-1999;	99US-0144352;	PR 14-OCT-1999;	99US-0159331;
PR 20-JUL-1999;	99US-0144632;	PR 14-OCT-1999;	99US-0159637;
PR 20-JUL-1999;	99US-0144884;	PR 14-OCT-1999;	99US-0159638;
PR 21-JUL-1999;	99US-0144814;	PR 18-OCT-1999;	99US-0159584;
PR 21-JUL-1999;	99US-0145086;	PR 21-OCT-1999;	99US-0160741;
PR 21-JUL-1999;	99US-0145088;	PR 21-OCT-1999;	99US-0160767;
PR 22-JUL-1999;	99US-0145085;	PR 21-OCT-1999;	99US-0160768;
PR 22-JUL-1999;	99US-0145087;	PR 21-OCT-1999;	99US-0160770;
PR 22-JUL-1999;	99US-0145089;	PR 21-OCT-1999;	99US-0160814;
PR 22-JUL-1999;	99US-0145192;	PR 21-OCT-1999;	99US-0160815;
PR 23-JUL-1999;	99US-0145145;	PR 22-OCT-1999;	99US-0160980;
PR 23-JUL-1999;	99US-0145218;	PR 22-OCT-1999;	99US-0160981;
PR 23-JUL-1999;	99US-0145224;	PR 22-OCT-1999;	99US-0160989;
PR 26-JUL-1999;	99US-0145276;	PR 25-OCT-1999;	99US-0161404;
PR 27-JUL-1999;	99US-0145913;	PR 25-OCT-1999;	99US-0161405;
PR 27-JUL-1999;	99US-0145918;	PR 25-OCT-1999;	99US-0161406;
PR 27-JUL-1999;	99US-0145919;	PR 26-OCT-1999;	99US-0161359;
PR 28-JUL-1999;	99US-0145951;	PR 26-OCT-1999;	99US-0161360;
PR 02-AUG-1999;	99US-0146386;	PR 26-OCT-1999;	99US-0161361;
PR 02-AUG-1999;	99US-0146388;	PR 28-OCT-1999;	99US-0161920;
PR 02-AUG-1999;	99US-0146389;	PR 28-OCT-1999;	99US-0161992;
PR 03-AUG-1999;	99US-0147038;	PR 28-OCT-1999;	99US-0161993;
PR 04-AUG-1999;	99US-0147204;	PR 29-OCT-1999;	99US-0162142;
PR 05-AUG-1999;	99US-0147302;		
PR 05-AUG-1999;	99US-0147192;		
PR 05-AUG-1999;	99US-0147260;		
PR 06-AUG-1999;	99US-0147303;		
PR 06-AUG-1999;	99US-0147416;		
PR 09-AUG-1999;	99US-0147493;		
PR 09-AUG-1999;	99US-0147935;		
PR 10-AUG-1999;	99US-0148171;		
PR 11-AUG-1999;	99US-0148319;		
PR 12-AUG-1999;	99US-0148341;		
PR 13-AUG-1999;	99US-0148565;		
PR 13-AUG-1999;	99US-0148684;		
PR 16-AUG-1999;	99US-0149368;		
PR 17-AUG-1999;	99US-0149175;		
PR 18-AUG-1999;	99US-0149426;		
PR 20-AUG-1999;	99US-0149722;		
PR 20-AUG-1999;	99US-0149723;		
PR 20-AUG-1999;	99US-0149929;		
PR 23-AUG-1999;	99US-0149902;		
PR 23-AUG-1999;	99US-0149930;		
PR 26-AUG-1999;	99US-0150566;		
PR 26-AUG-1999;	99US-0150884;		
PR 27-AUG-1999;	99US-0151065;		
PR 27-AUG-1999;	99US-0151066;		
PR 27-AUG-1999;	99US-0151080;		

Query Match

Best Local Similarity 79.9%; Score 1159; DB 21; Length 237;

Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSSRWNRITTYGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPRPPGYAFVFEPRDADD	60
Db	1	MSSRWNRITTYGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPRPPGYAFVFEPRDADD	60
QY	61	AIYGRDGDYDFGCRRLRVBIAGHGRFRFSPVDRYSSYSASAPRRSDYRVLTGLPESA	120
Db	61	AIYGRDGDYDFGCRRLRVBIAGHGRFRFSPVDRYSSYSASAPRRSDYRVLTGLPESA	120
QY	121	SWQDLKDHMRKAGDVCFSEVFPDRKMGVVDYSDNDMKYAIRKLDATFRNAPSSAYI	180
Db	121	SWQDLKDHMRKAGDVCFSEVFPDRKMGVVDYSDNDMKYAIRKLDATFRNAPSSAYI	180
QY	181	RVREYESRSVSRSPDDSKSYRSRSGRSPSCSYSSKSR	218
Db	181	RVREYESRSVSRSPDDSKSYRSRSGRSPSCSYSSKSR	218

RESULT 7

AAG09881

ID AAG09881 standard; Protein; 303 AA.

XX	AAAG09881;	PR	18-JUN-1999;	99US-0139462.
AC		PR	18-JUN-1999;	99US-0139463.
XX		PR	18-JUN-1999;	99US-0139750.
DT	17-OCT-2000 (first entry)	PR	18-JUN-1999;	99US-0139763.
XX		PR	21-JUN-1999;	99US-0139817.
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 7985.	PR	21-JUN-1999;	99US-0139899.
XX		PR	22-JUN-1999;	99US-0140353.
KW	Protein identification; signal transduction pathway; metabolic pathway;	PR	23-JUN-1999;	99US-0140354.
KW	hybridisation assay; Genetic mapping; gene expression control; promoter;	PR	23-JUN-1999;	99US-0140354.
KW	termination sequence.	PR	24-JUN-1999;	99US-0140895.
XX		PR	28-JUN-1999;	99US-0140823.
XX	Arabidopsis thaliana.	PR	29-JUN-1999;	99US-0140991.
XX		PR	30-JUN-1999;	99US-0141287.
PN	EP1033405-A2.	PR	01-JUL-1999;	99US-0141842.
XX		PR	01-JUL-1999;	99US-0142154.
PD	06-SEP-2000.	PR	02-JUL-1999;	99US-0142055.
XX		PR	06-JUL-1999;	99US-0142390.
PF		PR	08-JUL-1999;	99US-0142803.
XX		PR	09-JUL-1999;	99US-0142920.
XX	25-FEB-2000; 2000EP-0301439.	PR	12-JUL-1999;	99US-0142977.
XX		PR	13-JUL-1999;	99US-0143542.
PR	25-FEB-1999;	PR	14-JUL-1999;	99US-0143624.
PR	05-MAR-1999;	PR	15-JUL-1999;	99US-0144005.
PR	09-MAR-1999;	PR	16-JUL-1999;	99US-0144085.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144086.
PR	25-MAR-1999;	PR	19-JUL-1999;	99US-0144325.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144331.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144332.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144333.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144334.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144335.
PR	19-APR-1999;	PR	20-JUL-1999;	99US-0144352.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144884.
PR	23-APR-1999;	PR	21-JUL-1999;	99US-0144814.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145089.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148585.
PR	17-JUN-1999;	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930.

PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	23-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	07-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	13-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	28-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144086.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144684.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0145986.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161820.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161992.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147204.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147302.	PR	29-OCT-1999;	99US-0162142.
PR	05-AUG-1999;	99US-0147192.			
PR	05-AUG-1999;	99US-0147260.			
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			

Query Match 59.5%; Score 863; DB 21; Length 307;
Best Local Similarity 62.2%; Fred. No. 8.3e-77;
Matches 184; Conservative 33; Mismatches 45; Indels 34; Gaps 7;

Qy 1 MSSRWRTIYVGNLPGDIRKCEVEDLFYKYGPIVDLKIIPRPPGVAFVEFEDPRDAD 60
Db 1 MSSRSRTIYVGNLPGDIRKCEVEDLFYKYGPIVDLKIIPRPPGVAFVEFEDPRDAD 60
Qy 61 AIXGRDGYDFDGCRLRVEIAHGGRFSPSYDRYSSYS-----ASRAP 103

Db 61 AIHGRDGYDFDCHRLRVELAHGRR---SSDTRGSENGGRRGGRRGGRRGGRRGGRRGGRRGG 117

Qy 104 SRRSDYRLVTGLPPSSAQWDLKHMEKAGDVCFSEVFPDRKMGSGVVDYNNYDDMKYAL 163

Db 118 SRRSEFRVLTGLPSSASQWDLKHMEKGGDVCFSQVYRDARTGTGVVDYTCYEDMKYAL 177

Qy 164 KLDATFRNAPFSSAYIRVREYESSRSVSRSPDDSKSY-RSRSRSGPSCSYSSKSRSPV 222

Db 178 KCLDTEFRNAPFNGYVRVREYDSRKDSRSPRGESYSKSRSGRVS---SRSRSRSR 234

Qy 223 ARSTSPSRPLSRGRSLYSSVSRSGSLLRAGDWISQSRKSKS--RSRSGNSPVSVPV 276

Db 235 SRRSPKAK--SSRRSPAKSTSRSPGPR-----SKSRSPSPRRSRSRSPPLSPV 282

RESULT 9

AAB11417

ID AAB11417 standard; Protein; 303 AA.

AC AAB11417;

XX

XX

DT 23-FEB-2001 (first entry)

XX

DE A. thaliana SRP34/SRI protein.

XX

KW SR protein; splice-factor activity; plant; developmental behavior;

XX

KW flowering; crop plant; cereal; bean; rice; fruit.

XX

OS Arabidopsis thaliana.

XX

PN WO200605059-A1.

XX

XX

PD 02-NOV-2000.

XX

PF 20-APR-2000; 2000WO-AT00100.

XX

PR 23-APR-1999; 59AT-0000727.

XX

PA (OSTP) OSTERR FORSCH SETBERSDORF.

XX

PI Barta A, Lopato S, Kalyna M, Dornier S;

XX

DR WPI; 2000-687349/67.

XX

PT Novel proteins with splice-factor activity in plants, useful e.g. for

PT altering flowering time or development, and the nucleic acid that

PT encodes it -

XX

PS Disclosure; Fig 2; 67pp; German.

XX

CC This invention describes a novel protein (I) with splice-factor activity

CC in plants (I) modifies the choice of splice sites in many plant

CC pre-mRNAs. (I) (also the nucleic acid that encodes them and related

CC vectors or expression systems) are used: (i) to alter splice patterns in

CC plants, or their parts; (ii) to alter developmental behavior of plants;

CC and/or (iii) to delay flowering, particularly by at least 25% relative

CC to the wild type, especially in crop plants such as cereals, beans, rice

CC and fruit.

XX

SQ Sequence 303 AA;

Query Match 57.6%; Score 836; DB 21; Length 303;

Best Local Similarity 62.0%; Pred. No. 3.8e-74;

Matches 181; Conservative 32; Mismatches 49; Indels 30; Gaps 7;

Qy 1 MSSRNRTIYGNLPGDIRKEVEDLFVYKGPVIDIDIKIPRPPGYAFVEFEDPRADD 60

Db 1 MSSRSRTIYGNLPGDIRKEVEDLFVYKGPVIDIDIKIPRPPGYAFVEFEDPRADD 60

Qy 61 AIYGRDGYDFDCHRLRVELAHGRR---SSDTRGSENGGRRGGRRGGRRGGRRGGRRGGRRGG 107

Db 61 AIHGRDGYDFDCHRLRVELAHGRR---SSDTRGSENGGRRGGRRGGRRGGRRGGRRGGRRGG 117

Qy 108 DYRLVTLGLPPSSAQWDLKHMEKAGDVCFSEVFPDRKMGSGVVDYNNYDDMKYAL 167

Db 118 EERVLVTWLASSASQWDLKHMEKGGDVCFSQVYRDARTGTGVVDYTCYEDMKYAL 177

Qy 168 ATEFRNAPFSSAYIRVREYESSRSVSRSPDDSKSY-RSRSRSGPSCSYSSKSRSPV 226

Db 178 DTEFRNAPFNGYVRVREYDSRKDSRSPRGESYSKSRSGRVS---SRSRSRSRSR 234

Qy 227 SPRSRPLSRGRSLYSSVSRSGSLLRAGDWISQSRKSKS--RSRSGNSPVSVPV 276

Db 235 SPKAK--SSRRSPAKSTSRSPGPR-----SKSRSPSPRRSRSRSPPLSPV 278

RESULT 10

AAG15601

ID AAG15601 standard; Protein; 296 AA.

XX

AC AAG15601;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 15917.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

XX

KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147152.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.

PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148119.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156438.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match 55.5%; Score 805; DB 21; Length 296;
Best Local Similarity 62.5%; Pred. No. 4.3e-71;
Matches 183; Conservative 27; Mismatches 53; Indels 30; Gaps 10;

Qy	1	MSRNRRTIYGNLPGDIRKCEVEDLFYKYGPIVDILKIPPPPGYAFVEEDPRDADD	60
Db	1	MSGRSRSIYGNLPGDIRKCEVEDLFYKYGPIVDILKIPPPPGYCFVEHSDAED	60
Qy	61	AIYGRDGYDFDCRURVETIAHGRRFSPSDRY-----SSSY-----SASRAPSR	106
Db	61	AIKGRDGYNLDCRURVETIAHGRRFSPSDRY-----SSSY-----SASRAPSR	119
Qy	107	SDYRVLVNLGPPSASQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSNVDDMKYAIRKL	166
Db	120	SEBRFVIVRLGPPSASQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSNVDDMKYAIRKL	179
Qy	167	DATEFRNAPSSAYIRVREYE-SRSVRSRSPDDSKSYRSRSGRSPSCSYS-SKSRSVSPAR	224
Db	180	DATEFRNAPSSAYIRVREYE-SRSVRSRSPDDSKSYRSRSGRSPSCSYS-SKSRSVSPAR	238
Qy	225	SI-SRPLSRSLSYSSVRSRSGLLRAGDWISQSRSRSRSRNSP	272
Db	239	DLKSFRSLSRSLSRSS--PSPDKKSPRA-----MSRSXSRSRSRSRSP	284
RESULT 11			
Qy	AAG47935		
Db	AAG47935	standard; Protein; 136 AA.	
Qy	18-OCT-2000	(first entry)	
Db	18-OCT-2000	(first entry)	
Qy	Arabidopsis thaliana	protein fragment SEQ ID NO: 60473.	
Db	Arabidopsis thaliana	protein identification; signal transduction pathway; metabolic pathway;	
Qy	Arabidopsis thaliana	hybridisation assay; genetic mapping; gene expression control; promoter;	
Db	Arabidopsis thaliana	termination sequence.	
Qy	Arabidopsis thaliana		
Db	Arabidopsis thaliana		
Qy	EP1033405-A2		
Db	EP1033405-A2		
Qy	06-SEP-2000		
Db	06-SEP-2000		
Qy	25-FEB-2000	2000EP-0301439.	
Db	25-FEB-2000	2000EP-0301439.	
Qy	25-FEB-1999	99US-0121825.	
Db	25-FEB-1999	99US-0121825.	
Qy	03-MAR-1999	99US-0123180.	
Db	03-MAR-1999	99US-0123180.	
Qy	23-MAR-1999	99US-0125788.	
Db	23-MAR-1999	99US-0125788.	
Qy	25-MAR-1999	99US-0126264.	
Db	25-MAR-1999	99US-0126264.	
Qy	29-MAR-1999	99US-0126785.	
Db	29-MAR-1999	99US-0126785.	
Qy	01-APR-1999	99US-0127462.	
Db	01-APR-1999	99US-0127462.	
Qy	06-APR-1999	99US-0128234.	
Db	06-APR-1999	99US-0128234.	
Qy	08-APR-1999	99US-0128714.	
Db	08-APR-1999	99US-0128714.	
Qy	16-APR-1999	99US-0129845.	
Db	16-APR-1999	99US-0129845.	
Qy	19-APR-1999	99US-0130077.	
Db	19-APR-1999	99US-0130077.	
Qy	21-APR-1999	99US-0130449.	
Db	21-APR-1999	99US-0130449.	
Qy	23-APR-1999	99US-0130510.	
Db	23-APR-1999	99US-0130510.	
Qy	23-APR-1999	99US-0130891.	
Db	23-APR-1999	99US-0130891.	
Qy	28-APR-1999	99US-0131449.	
Db	28-APR-1999	99US-0131449.	
Qy	30-APR-1999	99US-0132048.	
Db	30-APR-1999	99US-0132048.	
Qy	30-APR-1999	99US-0132407.	
Db	30-APR-1999	99US-0132407.	
Qy	04-MAY-1999	99US-0132484.	
Db	04-MAY-1999	99US-0132484.	
Qy	05-MAY-1999	99US-0132485.	
Db	05-MAY-1999	99US-0132485.	
Qy	06-MAY-1999	99US-0132486.	
Db	06-MAY-1999	99US-0132486.	
Qy	07-MAY-1999	99US-0132487.	
Db	07-MAY-1999	99US-0132487.	
Qy	11-MAY-1999	99US-0134256.	
Db	11-MAY-1999	99US-0134256.	
Qy	14-MAY-1999	99US-0134218.	
Db	14-MAY-1999	99US-0134218.	
Qy	14-MAY-1999	99US-0134219.	
Db	14-MAY-1999	99US-0134219.	
Qy	14-MAY-1999	99US-0134370.	
Db	14-MAY-1999	99US-0134370.	
Qy	18-MAY-1999	99US-0134768.	
Db	18-MAY-1999	99US-0134768.	
Qy	18-MAY-1999	99US-0134941.	
Db	18-MAY-1999	99US-0134941.	

```

PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149802.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 45.7%; Score 663.5; DB 21; Length 136;
Best Local Similarity 92.5%; Pred. No. 1.4e-57;
Matches 136; Conservative 0; Mismatches 0; Indels 11; Gaps 1;

Qy 129 MRKAGDVCFSEVFPDRKMGSGVVDYSGNYSKRSRSPARSIKRLDATEFRNAPSSAYIRVREYER 188
Db 1 MRKAGDVCFSEVFPDRKMGSGVVDYSGNYSKRSRSPARSIKRLDATEFRNAPSSAYIRVREYER 60

Qy 189 SVSRSPDDSKSYRSRSGRSGPSCSYSKRSRSPARSIKRLDATEFRNAPSSAYIRVREYER 248
Db 61 SVSRSPDDSKSYRSRSGRSGPSCSYSKRSRSPARSIKRLDATEFRNAPSSAYIRVREYER 60

Qy 249 LLRAGDWISQSRKSRSRSGRSGNSPVSP 275
Db 118 -----SQSRKSRSGRSGNSPVSP 136

RESULT 12
AAG23884
ID AAG23884 standard; Protein; 142 AA.
XX AC AAG23884;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 27354.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX EN EPI033405-A2.
XX FD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 23-APR-1999; 99US-0130891.
XX PR 28-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
XX PR 05-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 07-MAY-1999; 99US-0132487.
XX PR 11-MAY-1999; 99US-0132863.
XX PR 14-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-0134768.

```

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135639.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140655.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.


```
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0147038.
PR 03-AUG-1999; 99US-0147039.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147132.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147305.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149388.
PR 18-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158233.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 44.6%; Score 647; DB 21; Length 128;
Best Local Similarity 100.0%; Fred. No. 5.6e-56; Mismatches 0; Indels 0; Gaps 0;
Matches 128; Conservative 0;

QY 129 MKKAGDVCFSEVFPDRKGMGVWVDSNYDDMKYAIRKLDATFKNFSAIRVREYESR 188
    |||||
Db 1 MKKAGDVCFSEVFPDRKGMGVWVDSNYDDMKYAIRKLDATFKNFSAIRVREYESR 60
    |||||

QY 189 SVSRSPDDSKYRSRSGRGPSCSYSSKSRVSPARSISPSRPLSRSRSLYSSVSRSGS 248
    |||||
Db 61 SVSRSPDDSKYRSRSGRGPSCSYSSKSRVSPARSISPSRPLSRSRSLYSSVSRSGS 120
    |||||

QY 249 LLRAGDWI 256
    |||||
Db 121 LLRAGDWI 128

RESULT 14
AAG30418
ID AAG30418 standard; Protein; 294 AA.
XX
AC AAG30418;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 36362.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
```


PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148584.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 01-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 44.6%; Score 646.5; DB 21; Length 294;
 Best Local Similarity 69.7%; Pred. No. 2.1e-55;
 Matches 124; Conservative 16; Mismatches 21; Indels 17; Gaps 2;

QY 1 MSSRWRTIYVGNLPGDIRKEVEDLFYKGPVVDIDLKIPPPGCGYAFVEFEDPRDADD 60
 Db 1 MSSRSRTIYVGNLPGDIRKEVEDLFYKGPVVDIDLKIPPPGCGYAFVEFEDARADD 60
 QY 61 AIYGRDGYDFDGLRLRVEIAHGGRRFSPVDRYSSGYSA-----SRAPSR 106
 Db 61 AIYGRDGYDFDGLRLRVEIAHGGRRFSPVDRYSSGYSA-----SRAPSR 117
 QY 107 SDYRVLVTGLPPSASQDLKDHVRKAGDVCFSVFPDRKMGSGVVDYSDMKYAIR 164
 Db 118 SEYRVVVSGLPSSASQDLKDHVRKAGDVCFSVFPDRKMGSGVVDYSDMKYAVR 175

RESULT 15
 AAB11418
 ID AAB11418 standard; Protein; 248 AA.
 AC AAB11418;
 XX
 DT 23-FEB-2001 (first entry)
 XX
 DE Human SF2/ASF protein.
 XX
 XX SR protein; splice-factor activity; plant; developmental behavior;
 XX flowering; crop plant; cereal; bean; rice; fruit.
 XX Homo sapiens.
 XX WO200065059-A1.
 XX 02-NOV-2000.
 XX 20-APR-2000; 2000WO-AT00100.
 XX 23-APR-1999; 99AT-0000727.
 XX (OSTP) OESTERR FORSCH SEIBERSDORF.
 XX Barta A, Lopato S, Kalyana M, Dörner S;
 XX WPI; 2000-687343/67.

Novel proteins with splice-factor activity in plants, useful e.g. for altering flowering time or development, and the nucleic acid that encodes it -
 Disclosure; Fig 2; 67pp; German.
 This invention describes a novel protein (I) with splice-factor activity in plants (I) modifies the choice of splice sites in many plant pre-mRNAs. (I) (also the nucleic acid that encodes them and related vectors or expression systems) are used: (i) to alter splice patterns in plants, or their parts; (ii) to alter developmental behavior of plants; and/or (iii) to delay flowering, particularly by at least 25% relative to the wild type, especially in crop plants such as cereals, beans, rice and fruit.

SQ Sequence 248 AA;

Query Match 40.1%; Score 582.5; DB 21; Length 248;

Best Local Similarity 53.6%; Pred. No. 3.5e-49;
 Matches 134; Conservative 28; Mismatches 47; Indels 41; Gaps 8;

QY 9 IYVGNLPGDIRKEVEDLFYKGPVVDIDLKIPPPGCGYAFVEFEDPRDADDALYGRDGY 68
 Db 18 IYVGNLPGDIRKEVEDLFYKGPVVDIDLKIPPPGCGYAFVEFEDPRDADDALYGRDGY 77
 QY 69 DFDGCRRLRVEIAHGGRR-----RFSFSDRYSSSYSSASRAPSRSDYRV 111
 Db 78 DYDGYRLRVEFPPRSGRGTRGGGGGGGAPRGRIGP-----PSRRSENRV 123
 QY 112 LVTGLPPSASQDLKDHVRKAGDVCFSVFPDRKMGSGVVDYSDMKYAIRKLDATF 171
 Db 124 VVSGLPSSGSGQDLKDHVRKAGDVCYADVRDG---TGVEFVRKEDMTYAVRKLDTKF 180
 QY 172 R-NAFSSAYIRVREYERSVSRSPDDSKSYRSRSGRSGPCSSYSSKSRSVSPARSI-SPR 229
 Db 181 RSHEGETAYIRVKVDGPRSPSYGRSGRSR-RSRSGRSRS---NSRSGSYSPRRSGSPR 236
 QY 230 SRPL-SRSGRS 238
 Db 237 YSPRHSRSGRS 246

Search completed: January 29, 2004, 02:50:13
 Job time : 51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 02:50:45 ; Search time 28 Seconds
(without alignments)
958.252 Million cell updates/sec

Title: US-10-014-927-19

Perfect score: 1451

Sequence: 1 MSSRWRTIYVGNLPGDIRK.....RSKSRSRNSPVPVWISG 279

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1159	79.9	237	2 E86223	hypothetical prote
2	865	59.6	303	2 F86158	alternative splici
3	839	57.8	303	2 S71185	splicing factor SF
4	646.5	44.6	294	2 T01307	alternative splici
5	597.5	41.2	243	2 T46221	PRE-MRNA SPLICING
6	582.5	40.1	248	2 A40040	alternative splici
7	517	35.6	221	2 S53075	splicing factor SR
8	498.5	34.4	201	2 C40040	alternative splici
9	497	34.3	201	2 S26404	alternative splici
10	497	34.3	292	2 B40040	alternative splici
11	435.5	30.0	344	2 S59043	splicing factor SR
12	432.5	29.8	350	2 A40459	nuclear phosphopro
13	431	29.7	494	2 A48133	pre-mRNA splicing
14	418.5	28.8	272	2 S53042	splicing factor SR
15	414	28.5	374	2 A37282	52K active chromatin
16	409	28.2	269	2 B47112	growth response pr
17	315.5	21.7	281	2 T26084	hypothetical prote
18	297	20.5	238	2 A57198	splicing factor, a
19	294	20.3	312	2 T26085	hypothetical prote
20	280	19.3	361	2 T42525	splicing factor-li
21	280	19.3	365	2 T37730	probable pre-mRNA
22	237.5	16.4	221	2 B42701	PR264 protein - ch
23	235	16.2	208	2 T15953	hypothetical prote
24	234.5	16.2	221	2 A42701	splicing factor SF
25	234	16.1	164	2 I54089	pre-mRNA splicing
26	234	16.1	164	2 S14016	X16 protein - mous
27	233	16.1	179	2 T34145	hypothetical prote
28	216	14.9	287	2 T50647	serine/arginine-ri
29	211	14.5	349	2 F85294	splicing factor At

30	208.5	14.4	302	2 T45890	splicing factor-li
31	205	14.1	350	2 T05797	splicing factor SR
32	204.5	14.1	284	2 T51304	splicing factor RS
33	198.5	13.7	249	2 E84791	hypothetical prote
34	193.5	13.3	271	2 T47978	splicing factor RS
35	189	13.0	135	2 A46398	RNA-binding protei
36	188	13.0	200	2 T05112	splicing factor 9G
37	187	12.9	200	2 T52627	splicing factor RS
38	186	12.8	274	2 A55335	myelin regulatory
39	185.5	12.8	286	2 T09704	probable arginine/
40	181.5	12.5	250	2 A84905	probable arginine/
41	181	12.5	196	2 E84638	probable R32p22 sp
42	173	11.9	196	2 T15917	hypothetical prote
43	169.5	11.7	414	2 JN0866	nucleolar protein
44	168.5	11.6	309	2 T47685	probable RNA bindi
45	163.5	11.3	275	2 T48696	mRNA splicing fact

ALIGNMENTS

RESULT 1

E86223

hypothetical protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002

C:Accession: E86223

R:Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonzo, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E86223

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-237 <STO>

A:Cross-references: GB:AE005172; NID:g3249109; PIDN:AAC24092.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein

Query Match 79.9%; Score 1159; DB 2; Length 237;

Best Local Similarity 100.0%; Pred. No. 8.3e-87;

Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSSRWRTIYVGNLPGDIRKCEVEDLFYKYGPIVDLKLIPPPGYAFVEFEDPDADD 60

Db 1 MSSRWRTIYVGNLPGDIRKCEVEDLFYKYGPIVDLKLIPPPGYAFVEFEDPDADD 60

Qy 61 AIYGRDGYDFDGRCLVEIAHGGRSPSPVDYSSYSASRAPSRSDYRVLVTGLPPSA 120

Db 61 AIYGRDGYDFDGRCLVEIAHGGRSPSPVDYSSYSASRAPSRSDYRVLVTGLPPSA 120

Qy 121 SWDLDKDHMKACGVCFSEVPDRKMGVVDYNSYDDMKYAIRKLDATFEFRNAPSAYI 180

Db 121 SWDLDKDHMKACGVCFSEVPDRKMGVVDYNSYDDMKYAIRKLDATFEFRNAPSAYI 180

Qy 181 RVREYSRSVSRSPDDSKYSRSTRSRGSPSCSYSSKSR 218

Db 181 RVREYSRSVSRSPDDSKYSRSTRSRGSPSCSYSSKSR 218

RESULT 2

F86158

alternative splicing factor SF2a [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002

A; Experimental source: cultivar Columbia; BAC clone T9C5
C; Genetics:
A; Map position: 3
A; Introns: 29/3; 51/3; 76/3; 124/1; 139/3; 159/1; 177/1;
A; Note: T9C5.30

Query Match	41.2%	Score 597.5;	DB 2;	Length 243;
Best Local Similarity	66.1%;	Pred. No. 3.3e-41;		
Matches 119;	Conservative	17;	Mismatches 29;	Indels 15; Gaps 3

[illegible]

Qy 61 AYGRDGVDFDGCRLRVETIAHGGRFPSPVDY-----SSSY-----SASRAPRR 106
 Db 61 AYKRGDGNLDGRLRVLAHGGRGQSSS-DRRGYGGGSGYCGGGGGGSAFPGVSRH 119

Oy		107	SYRVLVTGLPSSASQDLKDHNRKAGDVCFSEVFPDRKGMSGVVDYSNNYDDMKYAIRKL	166
	:	:	:	:
Dd		120	SSEFVIYRGLPSSASQDLKDHNRKAGDVCFEAETVRDSFGTYGVVDYTNYYDDMKYAVRNL	179

RESULT 6
A40040

C;Species: Homo sapiens (man)
C;Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 21-Jul-2000
C;Accession: A40040; B40041

Cell 66, 373-382, 1991
A;Title: Primary structure of the human splicing factor ASF reveals similarities
A;Reference number: A40040; MUID:91309149; PMID:1855257

A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-248 <GEA>

R; Krainer, A.R.; Mayeda, A.; Kozak, D.; Binns, G.
Cell 66, 383-394, 1991

A;Accession: B40041
A;Status: preliminary
A;Molecule type: mRNA

A; Cross-references: GB:M69040; NID:G338046; PID:AAA03476.1; PID:G338047
C; Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonuc
F: 100-100; Domain: ribonucleoprotein repeat homology <RRM1>

Query Match	40.1%;	Score 582.5;	DB 2;	Length 248;
Best Local Similarity	53.6%;	Pred. No. 5.6e-40;		

[illegible]

Qy	69	DFDGCRLRVEIAHGGR-----RFSPSVDRYSSYSASRAPRRSDRV	111
		: : :	
Db	78	DVDCGRIIVSRVPRSCGVCPCGCCCGCCCGCAGACDGGCVCDP	123
		: : :	

```

Qy 112 LVTGLPSASWDLKDHNKAGDVCFSEVFPDRKMGSVVDYSNYDDMKYAIRKLDATF 171
    : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Dh 124 VYSGI PPSGCSWDLKHNKPEAGCNVCA NVLPDG--TCGVYFVYFKDMVMVAVSYDNKTF 180
    : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```
Qy      172 R-NAFSAAYIRVREYERSSVRSPDDSKSYRSTRSRRCSCSVSKSRVSVPASRI - SPR 229  
          :       : | | | | | : | | | | | : | | | | |  
Dp      181 RSHEGEATYIRVKVDGPPSPSGRRSRS-RSPSPRSRS---NSRSSGVSPPSRGSPR 236
```

QY 230 SRFL-SRERS 238

Db 237 YSPRHSRS 246

RESULT 7
S59075

splicing factor SRp30c - human
C/Species: Homo sapiens (man)
C/Date: 15-Feb-1996 #sequence...
C/Accession: S59075

R; Screaton, G.R.; Cacares, J.F.; Mayeda, A.; Bell, M.V.; Flebanski, M.; Jackson, D.G.; B. EMBL J. 14, 4336-4349, 1995
A: Title: Identification and characterization of three members of the human SR family of
A: Reference number: S9042; UID: 96016206; PMID: 7558075

A;Accession: S59075
A;Status: preliminary
A;Molecule type: mRNA

A;Residues: 1-221 <SCr>
A;Cross-references: EMBL:U30835; NID:g1049077; PIDN:AA93369.1; PID:g1049078
C;Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein repeat-containing proteins
F;15-79/Domain: ribonucleoprotein repeat homology <RPM1>
F;112-172/Domain: ribonucleoprotein repeat homology <RM2>

Query Match 35.6%; Score 517; DB 2; Length 221;
Best Local Similarity 52.5%; Pred. No. 1e-34;
Matches 115; Conservative 30; Mismatches 38; Indels

Qy	9	IYVGNLPGDIRKCEVEDLFYKYGP VDILK-----IPRPDGYAFVEFEDPRDADDAY	63
Db	16	TVYGVNTDPNVSPVNDIENIEVKVGSTPFTETKNPDHGIIVD-----EAEVDESDDDADADAY	70

Qy 64 GRDGYDFGCLRVEI--AHGGRFFSVDRYSSSYASRAPRRSDRYRLVTGLPPSAS 121
| | : | | | | | : | | | | : | | | | |
Dh 71 GRNGVNYGCRIRVREPRDTYGCGWDECCP-----NGDPNRSDFRVI.VSGIPPSGS 123
| | : | | | | | : | | | | : | | | | |

```

122 WDLZDMRKAGDVCFSEVFPDRKGMSGVDSYNDMKYAIRKLDATFR-NAFSSAYI 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
124 WDLZDMRFBAGDVCFADVQKGVGM--VEVIRKENMEVAIRKIDDTFRSHEGTSYI 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

QY      181 RVREYERSVSRSPDDSKV---RSRSRSGPSCSYSK 216
      ||      |: || || || || || || || || || ||
Db      181 RV-----YPRESTSGVSRSCSSCSRGSDSPVSR 209

```

RESULT 8
C40040

C40040
alternative splicing factor ASF-3 - human
C;Species: Homo sapiens (man)
C;Date: 24-Jan-1992 #sequence_revision 24
C;Accession: C40040

R. Ge, H.; Zuo, P.; Manley, J. L.
 Cell 66, 373-382, 1991
 Title: Primary structure of the human splicing factor ASF reveals similarities with Drosophila ASF
 Reference number: A40040; MUID: 91309149; PMID: 1855257
 Accession: C40040

A;Accession: C40040
A;Status: preliminary
A;Molecule type: mRNA

A:Residues: 1-201 <GEA>
A:Cross-references: GB:M72709
C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein
F:17-81/Domain: ribonucleoprotein repeat homology <RRM1>
F:122-182/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match	34.4%;	Score 498.5;	DB 2;	Length 201;
Best Local Similarity	50.5%;	Pred. No. 2.8e-33;		
Matches 102;	Conservative	27;	Mismatches 38;	Indels 35

QY 9 IYVGNLPGDIRKCEVEDLFYKYGPIVIDIDLKIPRPPGYAFVEFEPEDDADDAIYGRDCY 68
||||| |||| :||:||||| ||||| :||||| ||||| :|||||
||||| |||| :||:||||| ||||| :||||| ||||| :|||||

```

QY      69  DFDCGLRLVETIAHGGR-----RFPSPVDRISSSYASRAPRRSDRYV 111
          ||:| | | | | | | | | | | | | | | | | | | | | | | | | |
          ||:| | | | | | | | | | | | | | | | | | | | | | | | | |

```

```
Db      78 DYDGYRLRVFFPSRGTRGGGSGGGAPGRYGP-----PSSRSENV 123
Qy     112 LVTLGPSASWQLDKDHMRKAGVCFSEVPFDPKRGMGVVDYSNYDDMKYAIRKLDATEF 171
       :|||||:||||||:||||||:||||||:||||||:||||||:||||||:||||||:
Db     124 VVSGLPPSGSWQLDKHMRKAGDVCVADVVRDG---TGWVEFVKEDMTYAVRKLDNTKF 180
Qy     172 RNAFPSSAYIRVREYESRVSRS 193
Db     181 R-SHEVGYYTRLFFDQNWIQWS 201

RESULT 9
S26404
alternative splicing factor ASF - mouse
C;Species: Mus musculus (house mouse)
C;Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 03-Dec-1999
C;Accession: S26404
R;tacke, R.; Boned, A.; Goridis, C.
Nucleic Acids Res. 20, 5482, 1992
A;Title: ASF alternative transcripts are highly conserved between mouse and man.
A;Reference number: S26404; MUID:93065226; PMID:1437571
A;Accession: S26404
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-201 <TAC>
A;Cross-references: EMBL:X66091
A;Note: The nucleotide sequence was submitted to the EMBL Data Library, May 1992
C;Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoproteins
F;17-81/Domain: ribonucleoprotein repeat homology <RRM1>
F;122-182/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match          34.3%; Score 497; DB 2; Length 201;
Best Local Similarity 54.4%; Pred. No. 3..7e-33;
Matches 99; Conservative 23; Mismatches 26; Indels 34; Gaps 3;

Qy     9 IYGNVLPGDIRKEVEDLFYKGPIVDIDLKIPRPFGYAFVFEPFDPRADDAIYGRDGY 68
       |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db    18 IYGNLPDPDIRTKIEDLVFYKAIRDIDLKNRRGGPPFAFVFEPFDPRADAIVYGRDGY 77
Qy    69 DFQGCRLRVEIAHGRR-----RFSPSVDRYSSSYSAAPSRRSDYEV 111
       :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db    78 DYDGYRLRVFFPSRGTRGGGSGGGAPGRYGP-----PSSRSENV 123

Qy     112 LVTLGPSASWQLDKDHMRKAGVCFSEVPFDPKRGMGVVDYSNYDDMKYAIRKLDATEF 171
       :|||||:||||||:||||||:||||||:||||||:||||||:||||||:||||||:
Db     124 VVSGLPPSGSWQLDKHMRKAGDVCVADVVRDG---TGWVEFVKEDMTYAVRKLDNTKF 180
Qy     172 RN 173
Db     181 RS 182

RESULT 10
B40040
alternative splicing factor ASF-2 - human
C;Species: Homo sapiens (man)
C;Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 21-Jul-2000
C;Accession: B40040
R;Ge, H.; Zuo, P.; Manley, J.L.
Cell 66, 373-382, 1991
A;Title: Primary structure of the human splicing factor ASF reveals similarities with Dn
A;Reference number: A40040; MUID:91309149; PMID:1855257
A;Accession: B40040
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-292 <GEA>
A;Cross-references: GB:M72709; NID:g179073; PID:AAA35564.1; PID:g179074
C;Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoproteins
F;17-81/Domain: ribonucleoprotein repeat homology <RRM1>
F;122-182/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match          34.3%; Score 497; DB 2; Length 292;
Best Local Similarity 54.4%; Pred. No. 6e-33;
```

Matches 99; Conservative 23; Mismatches 26; Indels 34; Gaps 3;

QY 9 IYVGNLPGDIRKEVEDLFYKGPVIDLKIIPRPPPGYAFVEFEDPRDADAIYGRDGY 68
| | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18 IYVGNLPPIRKVDIEDFYKGAITDIDLKVRGGPFPFAFVEFEDPRDAEAVYGRDGY 77
| | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 69 DFGCLRLVEIAHGG-----RFSPSVDRYSYSSASRASRSRDYRV 111
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 78 DYDGRLRVEFRSGRGTCRGCGGGGGGAGPRGYVP-----PSRSSEN RV 123
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 112 LVTGLPPSASWDLKHMRKAGVCSEYFPDPKMGSGVVDSNYDDMKYAIRKLDATEF 171
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 124 VWSGLPSPGWDLKHMRKAGVCYADVIRDG---TGVVEFRKEDMTYVRKLONTKF 180
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 172 RN 173
| : |
Db 181 RS 182

RESULT 11
S59043
Spilling factor SRP55 - human
C/Species: Homo sapiens (man)
C/Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 28-May-1999
C/Accession: S59043
EMBO J. 14, 4336-4349, 1995
R/Sreston, G.R.; Caceres, J.F.; Mayeda, A.; Bell, M.V.; Piebanski, M.; Jackson, D.G.; E
A/Title: Identification and characterization of three members of the human SR family of
A/Reference number: S59042; PMID:96016206; PMID:7556075
A/Accession: S59043
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-344 <GB>
A/Cross-references: GB:U30883; NID:G1049087; PIDN:AAA93073.1; PID:G1049088
C/Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotei
F/3-62/Domain: ribonucleoprotein repeat homology <RRM1>
F/111-173/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 30.0%; Score 435.5; DB 2; Length 344;
Best Local Similarity 38.8%; Pred. No. 7.3e-28;
Matches 123; Conservative 44; Mismatches 89; Indels 61; Gaps 12;

QY 9 IYVGNLPGDIRKEVEDLFYKGPVIDLKIIPRPPPGYAFVEFEDPRDADAIYGRDGY 68
| : | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4 VTIGRLSYNVREKDIORFFSGYGRLLVDLK----NGYGFVEFEDSRDADAVYELNGK 58
| : | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 69 DFGCLRLVEIAHGGRRFPSVDYIS-----SSVSASRAFSR-----RSDYRVLYT 114
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 59 ELCEGHVIEHAAGRP--DRDGSYSGRSGGGYSRRSTSGRDKYGPPVRYEYRLIVE 115
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 115 GLPPSASWDLKHMRKAGVCSEYFPDPKMGSGVVDSNYDDMKYAIRKLDATEF--- 171
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 116 NLSSRCSWDLCKDFMQAGEVTVADAHKRTN-EGVIEFESYSDMKRALDKLDGTINGR 174
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 172 -----RNAFSAIYRV-----EYESRSVSRSPDCSKYSRSRSRGP 209
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 175 NIRLIKDPTSHRRYSVGSRSRSRRRSRSRSSRSRSRSRSRSRSRSRSRSRSRSGKR 233
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 210 SCYSIS--KSRSVSPARSISR---SRPLSRERSLYSYVSRSGSLARA-----GDWISQ 258
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 234 SRSRSGKRSKSKSKPKSDRGSHSRSRSKOEYEK-SRSGRSRSPKENGKGDIKSK 292
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 259 SRSGRSRSGRSNSPVSP 275
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 293 SRSGRSRSGNSPLVPVP 309
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
A40459
nuclear phosphoprotein SRP55 - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 24-Sep-1999
C/Accession: A40459; S14620

R:Roth, M.B.; Zahler, A.M.; Stolk, J.A.
J. Cell Biol. 115, 587-596, 1991
A:Title: A conserved family of nuclear phosphoproteins localized to sites of polymerase
A:Reference number: A40459; MUID:92011900; PMID:1717489
A:Accession: A40459
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-350 <R0T>
A:Cross-references: EMBL:X58720; NID:g8496; PIDN:CAA41556.1; PID:g8497
C:Genetics:
A:Gene: SR55
A:Cross-references: FlyBase:FBgn0004587
C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein
C:Keywords: phosphoprotein; pre-mRNA splicing
F:5-64/Domain: ribonucleoprotein repeat homology <RRM3>
F:116-178/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 29.8%; Score 432.5; DB 2; Length 350;
Best Local Similarity 40.4%; Pred. No. 1.3e-27;
Matches 124; Conservative 40; Mismatches 92; Indels 51; Gaps 9;

Qy 9 IYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPPPPGYAFVEFDPDPRDADDAIYGRDGY 68
Db 6 VVVGILPYGVREDLIERFKYGRTRDILK-----NGYGFVEFEDYDADDAIYVINGK 60

Qy 69 DFDGCLRLVEIAHGRFPSPVDYSSY-----SASR-APRRSDYRVLV 113
Db 61 ELGLGVVVEPARGSGS-NRDYDDRYGRRGGGGRYNEKSSRYGPPLRTEYRLIV 119

Qy 114 TGLPSSAQDLKDHWRKAGDVCSEVPDRKMGSGVVDYNDYDMKYAIRKLDATPR- 172
Db 120 ENLSSRVSNQDLKYNRQAGEVYADAHKQRN-EGVVEFASLSDMKYAIRKLDTELNG 178

Qy 173 -----NAFSSAIRVREYSRSPDSDSKSYR---SRSRSGPSCSY 214
Db 179 RRIHVLVDRGRSGGGGGRGRSSSRSSSRSSSRSSSRSSSRSSSRSGRSKSRG 238

Qy 215 SKRSRVSPP-----ARISPRSRPLSRGSLYSSVSRGSLLRAGDWSQKRS-----KSR 264
Db 239 GRSKSKSPVKSRSRSGRSNKRSDYKSKSKSHSTRSPKGRSDSRSTRYSVKRSR 298

Qy 265 SRSRNS 271
Db 299 SRSRSGS 305

RESULT 13
A48133
pre-mRNA splicing SRP75 - human
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1995 #sequence_revision 30-Jun-1995 #text_change 05-Nov-1999
C:Accession: A48133
R:Zahler, A.M.; Neugebauer, K.M.; Stolk, J.A.; Roth, M.B.
Mol. Cell. Biol. 13, 4023-4028, 1993
A:Title: Human SR proteins and isolation of a cDNA encoding SRP75.
A:Reference number: A48133; MUID:93309435; PMID:8321209
A:Accession: A48133
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-494 <ZAH>
A:Cross-references: GB:L14076; NID:g307437; PIDN:AAA36649.1; PID:g307438
A:Note: parts of this sequence were confirmed by peptide sequencing
C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein
C:Keywords: phosphoprotein; pre-mRNA splicing
F:3-62/Domain: ribonucleoprotein repeat homology <RRM3>
F:105-167/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 29.7%; Score 431; DB 2; Length 494;
Best Local Similarity 40.1%; Pred. No. 2.7e-27;
Matches 124; Conservative 39; Mismatches 92; Indels 54; Gaps 10;

Qy 9 IYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPPPPGYAFVEFDPDPRDADDAIYGRDGY 68
Db 6 VVVGILPYGVREDLIERFKYGRTRDILK-----NGYGFVEFEDYDADDAIYVINGK 60

Db 4 VYIGRLSYQABERDVERFFKYGKILEVDLK-----NGYGFVEFDDLRDADDAIYVINGK 58
Qy 69 DFDGCLRLVEIAHGRFPSPVDYSSYASRA-----PSRSDYRVLVLTGLPPSASW 122
Db 59 DLGGERVIVEHARGPRR-DGSYSGRSGYGRSGRDKYGPPTREYRLIVENLSRCSW 117
Qy 123 QDLKDHWRKAGDVCSEVPDRKMGSGVVDYNDYDMKYAIRKLDATPR----- 172
Db 118 QDLKDYNRQAGEVYADAHKQRN-EGVIEFVSYDMKGALEKLDGTEVNGRKIRLVEDK 176
Qy 173 --NAFSSAIRVREYSRSPDSDSKSYR-----PDDSKSYRSGRSGPSCSY- SKRSRV 220
Db 177 PGRRRRSYRSH-SRERSRSHRSHKSRSGSGSKSHSKSRSGRSGRSGRSGR 235
Qy 221 SPARSISPR-----SRPLSRSLYSSVSRGSLLRAGDWI-----SOSRSK 262
Db 236 SQSKRSKKEKSRSPSKDKSRSGSHSAGSKSRSKDQAEKIQNDNVGPKPSRSHK 295
Qy 263 SRSRSGNS 271
Db 296 SKSKSRSGS 304

RESULT 14
S59042
splicing factor Srp40 - human
C:Species: Homo sapiens (man)
C>Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 28-May-1999
C:Accession: S59042
R:Streaton, G.R.; Caceres, J.F.; Mayeda, A.; Bell, M.V.; Plebanski, M.; Jackson, D.G.; B.
EMBO J. 14, 4336-4349, 1995
A:Title: Identification and characterization of three members of the human SR family of
A:Reference number: S59042; MUID:96016206; PMID:7556075
A:Accession: S59042
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-272 <SCR>
A:Cross-references: NID:g1049079; PIDN:AAA93070.1; PID:g1049080
C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein
F:5-64/Domain: ribonucleoprotein repeat homology <RRM3>
F:109-171/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 28.8%; Score 418.5; DB 2; Length 272;
Best Local Similarity 42.5%; Pred. No. 1.3e-26;
Matches 121; Conservative 29; Mismatches 90; Indels 45; Gaps 10;

Qy 9 IYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPPPPGYAFVEFDPDPRDADDAIYGRDGY 68
Db 6 VYIGRLNPAAREKDVVERFFKYGIRIDILK-----RGFGFVEFEDPRDADDAIYELDGK 60

Qy 69 DFDGCLRLVEIAH-----GGRFSPSPVDYSSYASR-----APSRSDYRVLVITG 115
Db 61 EL--CSRVTIEHARARSGRG-----GRGKYSRDRSRRPRNDRNAPPVVTENRLIVEN 114

Qy 116 LPPSAGNQDLKDHWRKAGDVCSEVPDRKMGSGVVDYNDYDMKYAIRKLDATPRNAP 175
Db 115 LSRVSNQDLKDFWRQAGEVTFADAHKPLN-EGVVEFASYGDLKNAIEKLSKEI-NGR 172

Qy 176 SSAIRVREYSRSPDSDSKSYRSGRSGPSCSYSKSRSGSPARSISPRSPLSR 235
Db 173 KIKLIEGSKHSHSRSGRSGRSGRSGRSGRSGRSGRSGRSGRSGRSGRSGRSGR 222

Qy 236 SRSLSVSRSG-----GSLLRAGDWISQSRSGRSGRSGRSGNS 271
Db 223 SRSKSRSGSPVPEKSKQKSGSSRSGSPASVDRQSRSGRSGRSGRSGRSGRSGRSGR 267

RESULT 15
A37282
52K active chromatin boundary protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 28-Feb-1992 #sequence_revision 28-feb-1992 #text_change 02-Jul-1998
C:Accession: A37282

Search completed: January 29, 2004, 03:02:27
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2004, 21:21:46 : Search time 18 Seconds
(without alignments)
728.914 Million cell updates/sec

Title: US-10-014-927-19
Perfect score: 1451
Sequence: 1 MSRRNRTIYVGNLPGDIRK.....RSKSRSRNSPVPVISG 279

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	865	59.6	303	1	SFR1_ARATH	Q22315 arabidopsis
2	582.5	40.1	247	1	SFR9_HUMAN	Q07955 homo sapien
3	517	35.6	221	1	SFR9_HUMAN	Q13247 homo sapien
4	440.5	30.4	344	1	SFR6_HUMAN	Q13247 homo sapien
5	431	29.7	494	1	SFR4_HUMAN	Q08170 homo sapien
6	426.5	29.4	489	1	SFR4_MOUSE	Q8ve97 mus musculus
7	421	29.0	375	1	SRS5_DROME	P26686 drosophila
8	418.5	28.8	272	1	SFR5_HUMAN	Q13243 homo sapien
9	409	28.2	269	1	SFR5_MOUSE	Q09167 rattus norv
10	374.5	25.8	270	1	SFR5_MOUSE	Q35326 mus musculus
11	297	20.5	238	1	SFR2_HUMAN	Q16629 homo sapien
12	280	19.3	365	1	SFR2_SCHPO	P78814 schizosacch
13	237.5	16.4	221	1	SFR2_CHICK	P30352 gallus gall
14	234.5	16.2	221	1	SFR2_MOUSE	Q62093 mus musculus
15	234	16.1	164	1	SFR3_HUMAN	P23152 homo sapien
16	232	16.0	208	1	YSX2_CAEEL	Q10021 caenorhabdi
17	226.5	15.6	221	1	SFR2_HUMAN	Q01130 homo sapien
18	210	14.5	356	1	RS41_ARATH	P92366 arabidopsis
19	207.5	14.3	250	1	RS31_ARATH	P92964 arabidopsis
20	205	14.1	350	1	RS40_ARATH	P92965 arabidopsis
21	189	13.0	135	1	RBPI_DROME	Q02427 drosophila
22	173	11.9	196	1	SFR2_CAEEL	Q09511 caenorhabdi
23	172	11.9	283	1	SFR2_MOUSE	Q60701 mus musculus
24	169.5	11.7	414	1	NOP3_YEAST	Q01560 saccharomyc
25	164.5	11.3	275	1	SRP1_SCHPO	Q10193 schizosacch
26	162	11.2	391	1	ROX1_HUMAN	P38159 homo sapien
27	161.5	11.1	197	1	ROX1_DROME	Q24491 drosophila
28	156.5	10.8	388	1	ROG_MOUSE	Q35479 mus musculus
29	154.5	10.6	437	1	RUI7_HUMAN	P08621 homo sapien
30	153	10.5	429	1	HRBI_YEAST	P38922 saccharomyc
31	152.5	10.5	378	1	RUI7_MOUSE	Q62376 mus musculus
32	150.5	10.4	629	1	PAB2_ARATH	P42731 arabidopsis
33	147	10.1	386	1	TIAL_MOUSE	P52912 mus musculus

34	146.5	10.1	288	1	TR2B_HUMAN	Q15815 homo sapien
35	145.5	10.0	471	1	RUI7_XENLA	P09406 xenopus lae
36	145	10.0	834	1	CWFM_SCHPO	Q9p639 schizosacch
37	144.5	10.0	386	1	TIAL_HUMAN	P31483 homo sapien
38	144	9.9	272	1	TRA2_DROVI	O02008 drosophila
39	142.5	9.8	632	1	PABP_DROME	P21187 drosophila
40	141.5	9.8	289	1	ROCI_ARATH	Q9uu44 arabidopsis
41	140.5	9.7	273	1	ROCI_NICSY	Q08935 nicotiana s
42	140.5	9.7	1089	1	Y555_HUMAN	Q9ukj3 homo sapien
43	139	9.6	264	1	TRA2_DROME	P19018 drosophila
44	138	9.5	282	1	ROC_XENLA	P19600 xenopus lae
45	137.5	9.5	427	1	GBP2_YEAST	P25555 saccharomyc

ALIGNMENTS

RESULT 1
SFR1_ARATH STANDARD; PRT; 303 AA.
AC Q22315; Q22314; Q39201;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Pre-mRNA splicing factor SFR2 (SRI protein).
GN SFR2 OR ATIG02840 OR F22D16.16 OR F22D16_30.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95372342; PubMed=7644475;
RA Lazar G., Schaal T., Maniatis T., Goodman H.M.;
RT "Identification of a plant serine-arginine-rich protein similar to the
RL mammalian splicing factor SFR2/ASF.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:7672-7676(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Su C.-L., Schuler W.A.;
RN Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA MEDLINE=21018719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
Dunn P., Etgu P., Feldlyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziali A.,
Mitscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana.";
RL Nature 408:816-820(2000).
CC -!- FUNCTION: CAN PROMOTE SPLICE SITE SELECTION IN VITRO PRESUMABLY BY
CC ANTAGONIZING THE EFFECTS OF THE A1 HETEROGENEOUS NUCLEAR
CC RIBONUCLEOPROTEIN. MAY HAVE AN ESSENTIAL FUNCTION DURING EARLY
CC PLANT DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;

CC Name=1; Synonyms=SF2A;
CC IsoId=022315-1; Sequence=Displayed;
CC Name=2; Synonyms=SF2B;
CC IsoId=022315-2; Sequence=VSP_005859, VSP_005860;
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES IN THE RS
CC DOMAIN.
CC -!- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.
CC -!- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M98340; AA32856.1; -;
CC EMBL; AF001035; AAB71385.1; -;
CC EMBL; AF001035; AAB71386.1; -;
CC EMBL; AC009525; AAF02881.1; -;
CC PIR; F86158; F86158.
CC PIR; S71185; S71185.
CC HSP; F19339; 1SXL.
CC InterPro; IPR000504; RNA_rec_mot.
CC SMART; SM0076; xrm; 2.
CC SMART; SM00360; RRM; 2.
CC PROSITE; PS0102; RRM; 2.
CC PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
CC RNA processing; mRNA splicing; Nuclear protein; RNA-binding; Repeat;
CC Phosphorylation; Alternative splicing.
CC DOMAIN 7 82 RNA-BINDING (RRM) 1.
CC FT DOMAIN 96 112 GLY-RICH (HINGE REGION).
CC FT DOMAIN 119 197 RNA-BINDING (RRM) 2.
CC FT DOMAIN 198 273 ARG/SER-RICH (RS DOMAIN).
CC FT DOMAIN 274 303 LYS/PRO/SER-RICH (PSK DOMAIN).
CC FT VARSPLIC 267 272 SRSRSR -> YGFTYD (in isoform 2).
CC FT VARSPLIC 273 303 Missing (in isoform 2).
CC FT VARSPLIC 273 303 Missing (in isoform 2).
CC FT CONFLICT 125 125 G -> W (IN REF. 1).
CC FT CONFLICT 127 127 P -> A (IN REF. 1).
CC FT CONFLICT 139 140 MR -> IA (IN REF. 1).
CC SQ SEQUENCE 303 AA; 33729 MW; 6530F9CB628B8EFA CRC64;

Query Match 59.6%; Score 865; DB 1; Length 303;
Best Local Similarity 63.0%; Pred. No. 1.3e-60;
Matches 184; Conservative 33; Mismatches 45; Indels 30; Gaps 7;

Qy 1 MSSRWRTIYVGNLPGDIRKCEVEDLFYKGPVIVDIDLKIPPPPGYAFVEFEDPRADD 60
Db 1 MSSRSRTIYVGNLPGDIREREVEDLFYKGPVIVDIDLKIPPPPGYAFVEFEDPRADD 60

Qy 61 AIYGRGDYDFGCRURVIAHGRFRFSVDVRYSSYS- - -SSDTRGSFNGGGRGGRGDSGSGPSRRS 107
Db 61 AIHGRGDYDFGHRURVIAHGRFR- - -SSDTRGSFNGGGRGGRGDSGSGPSRRS 117

Qy 108 DYRLVLTGLPSPASQDLKDMRKAGDVCSEVPFPRKMGSVVDYNSYDDMKYAIKLD 167
Db 118 EFRVLVTGLPSPASQDLKDMRKAGDVCSEVPFPRKMGSVVDYNSYDDMKYAIKLD 177

Qy 168 ATEFPAASSYAIRVEYESVSRSPDPSKSY-RSRSRSGPSCSVSSKRSRSPARSI 226
Db 178 DTEFPAASSYAIRVEYESVSRSPDPSKSY-RSRSRSGPSCSVSSKRSRSPARSI 234

Qy 227 SPSPRPLSRSLYSSVSRSGSLLRAGDWISQSRKS- -RSRSRNSGVPSPV 276
Db 235 SPKAK- -SRSPSPAKSTSRSGPR- - -SKSRSPSPRSPSRSPSRPLSPV 278

RESULT 2
SFRL_HUMAN

ID AC Q07955; Q13809;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Splicing factor, arginine/serine-rich 1 (pre-mRNA splicing factor SF2,
DE P33 subunit) [Alternative splicing factor ASF-1].
GN SFPS1 OR SF2P33 OR ASF OR SF2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_SEQUENCE FROM N.A., AND SEQUENCE OF 143-160 AND 166-174.
RP MEDLINE=91309150; PubMed=1830244;
RX Krainer A.R., Mayeda A., Korak D., Binns G.;
RA "Functional expression of cloned human splicing factor SF2: homology
RT to RNA-binding proteins, U1 70K, and Drosophila splicing
RT regulators";
RL Cell 66:383-394 (1991).
RN [2]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND ALTERNATIVE SPLICING.
RX MEDLINE=91309149; PubMed=1855257;
RA Ge H., Zuo P., Manley J.L.;
RT "Primary structure of the human splicing factor ASF reveals
RT similarities with Drosophila regulators";
RL Cell 66:373-382 (1991).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM ASF-1).
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loguercio N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4]
RP SEQUENCE OF 122-139.
RX MEDLINE=92249775; PubMed=1577277;
RA Zahler A.M., Lane W.S., Stolk J.A., Roth M.B.;
RT "SR proteins: a conserved family of pre-mRNA splicing factors";
RL Genes Dev. 6:837-847 (1992).
RN [5]
RP INTERACTIONS IN SPLICEOSOME ASSEMBLY.
RX MEDLINE=94084782; PubMed=8261509;
RA Wu J.Y., Maniatis T.;
RT "Specific interactions between proteins implicated in splice site
RT selection and regulated alternative splicing";
RL Cell 75:1061-1070 (1993).
RN [6]
RP FUNCTION IN RECRUITMENT OF U1-70K TO PRE-MRNA.
RX MEDLINE=94187841; PubMed=8139654;
RA Kohetz J.D., Jamison S.F., Will C.L., Zuo P., Luhrmann R.,
RA Garcia-Blanco M.A., Manley J.L.;
RT "Protein-protein interactions and 5'-splice-site recognition in
RT mammalian mRNA precursors";
RL Nature 368:119-124 (1994).
RN [7]


```

DR Pfam; PF00076; rtm; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS0102; RRM; 2.
DR PROSITE; PS0030; RRM RNP 1; FALSE NEG.
KW Nuclear protein; RNA-Binding; mRNA splicing; Repeat; Phosphorylation.
FT DOMAIN 14 89 RNA-BINDING (RRM) 1.
FT FT DOMAIN 91 100 GLY-RICH (HINGE REGION).
FT FT DOMAIN 111 187 RNA-BINDING (RRM) 2.
FT FT DOMAIN 188 200 ARG/SER-RICH (RS DOMAIN).
SQ SEQUENCE 221 AA; 25542 MW; 1EE7BD8601CD80C0 CIRC64;

Query Match 35.6%; Score 517; DB 1; Length 221;
Best Local Similarity 52.5%; Pred. No. 1.3e-33;
Matches 115; Conservative 30; Mismatches 38; Indels 36; Gaps 8

QY 9 IVVGNLPDGIKKVEEDLFYKYGPVIDIDLK-----IPRPQYAFVEFEDPRDADDAIY 63
DB 16 IVVGNLPDIDVREKLEDLFYKGIREFIELKNRHGLVP-----FAFVREDPRDAEDAIY 70
QY 64 GRDGYDFDGCRLRVEI--AHGGRFPSPVDYRYSYSSYASRAPRRDGYLVLTGLPPSAS 121
DB 71 GRNGDYDQCRLRVEFFRTYGGRGWPRGGR-----NGPRTRRSDFVLVSLGPPSGS 123
QY 122 WDLKDHNRKAGDYCFSEVPDRKSGVVDYNDYDMKYATRKLDATFR-NATSSAVI 180
DB 124 WDLKDHNRKAGDYCFADVKQGVGM--VEYLKEDMEYALRKLDITFRSHEGETSYI 180
QY 181 RVREYESRSVRSPPDSKSY---RSRSRSGPSCSYSSK 216
DB 181 RV-----YPERSTSYGYSRSRSGSRGRDSPYQSR 209

RESULT 4
SFR6 HUMAN
ID _SFR6 HUMAN STANDARD; PRT: 344 AA.
AC Q13247; Q13244; Q13245; Q96J06; Q9UJB8; Q9Y3N7;
DT 15-DEC-1998 (Rel. 37, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-FEB-2003 (Rel. 42, Last annotation update)
DE Splicing factor, arginine/serine-rich 6 (Pre-mRNA splicing factor
DE SFR55)
GN SFR36 OR SFR55.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxId=9606;
RN [1]_TaxId=9606;
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Colon;
RX MEDLINE=96016206; PubMed=7556075;
RA Sreaton G.G., Caceres J.F., Mayeda A., Bell M.V., Plebanski M.,
RA Jackson D.G., Bell J.I., Krainer A.R.;
RT "Identification and characterization of three members of the human SR
RT family of pre-mRNA splicing factors.";
RL EMBO J. 14:4336-4349(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Sabbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle B., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehaeslatho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLeay K., McMurray A.A.,

```

RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prithalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.L., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.W., Williams S.A.,
RA Whittenhead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilmink I., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM SRP55-2).
RR Laird G.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM SRP55-1).
RR TISSUE=Placenta;
RC MEDLINE=22398257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Haieh F.,
RA Datchenko L., Marudina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshituki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hate S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muszy D.M., Sodergren E.J., Lu X.J., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.J., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP SEQUENCE OF 21-27 AND 47-55.
RX MEDLINE=52249775; PubMed=1577277;
RA Zahler A.M., Lane W.S., Stolk J.A., Roth M.B.;
RT "SR proteins: a conserved family of pre-mRNA splicing factors.";
RL Genes Dev. 6:837-847(1992).
CC -1- FUNCTION: PLAYS A ROLE IN CONSTITUTIVE SPLICING AND CAN MODULATE
CC THE SELECTION OF ALTERNATIVE SPLICE SITES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=3;
CC Name=SRP55-1;
CC IsoId=Q13247-1; Sequence=Displayed;
CC Name=SRP55-2;
CC IsoId=Q13247-2; Sequence=VSP_005869, VSP_005870;
CC Name=SRP55-3;
CC IsoId=Q13247-3; Sequence=VSP_005871;
CC -1- PM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES IN THE RS
CC DOMAIN (BY SIMILARITY).
CC -1- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.
CC -1- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).

DR EMBL; U30883; AAA93073.1; --
DR EMBL; U30828; AAA93071.1; --
DR EMBL; U30829; AAA93072.1; --

```

ENML; AL031681; CAB43960.1; -.
DR ENML; AL031681; CAB43961.1; -.
DR ENML; BC006832; AAH06832.1; -.
DR PIR; S59043; S59043.
DR Genew; HGNC:10788; SFRS6.
DR GK; Q13247; -.
DR MIM; 601944; -.
DR GO; GO:0008248; F:pre-mRNA splicing factor activity; TAS.
DR GO; GO:0006376; P:RNA splice site selection; TAS.
DR InterPro; IPR000504; RNA_rec_mot.
DR PFAM; PF000076; rrm; 2.
DR SMART; SM00360; rrm; 2.
DR PROSITE; PS50102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW Nuclear protein; RNA-binding; mRNA splicing; Alternative splicing;
Repeat; Phosphorylation.
FT DOMAIN 1 72 RNA-BINDING (RRM) 1.
FT DOMAIN 87 90 GLY-RICH (HINGE REGION).
FT DOMAIN 110 183 RNA-BINDING (RRM) 2.
FT DOMAIN 184 343 ARG/SER-RICH (RS DOMAIN).
FT VARSPPLIC 86 135 SGGGYSRRSTRGRDKYGPVTEYELIVENSRRCSWQDL
KQFMROAGE -> MTNGAEAVSTEAKMTAFPDWFLFTLC
FT FT FT FT DCPMTLWLTLPKMTAAFC (in isoform
FT FT FT FT SRP55-2)
FT FT FT FT /FTID-VSP 005869.
FT FT FT FT Missing (in isoform SRP55-2).
FT FT FT FT /FTID-VSP 005870.
FT FT FT FT RVSPPPKATGRSRSRSSKSRSRSSSRD -> LKLGA
FT FT FT FT RMSQOGTSLXSLASSC (in isoform SRP55-3).
FT FT FT FT /FTID-VSP 005871.
FT FT FT FT R -> H (IN REF. 1).
FT FT FT FT R -> H (IN REF. 1).
SQ SEQUENCE 344 AA; 39586 MW; 72305506C5948B94 CRC64;

Query Match 30.4%; Score 440.5; DB 1; Length 344;
Best Local Similarity 39.1%; Pred. No. 2.2e-27;
Matches 124; Conservative 44; Mismatches 88; Indels 61; Gaps 12;

QY 9 IVGNLPGDIRKEVEDLPYKGPVIDLKIPLPPPGYAFVEFEDPRDADDAIYGRDGY 68
Db 4 VVIGRLSYNVREKDIQRFSSGYGRLEVDLK-----NGYGFVEFEDSRDADDAVYELNGK 58

QY 69 DFEGCELRVEIAHGGRFPSPVDIYS-----SSYSASRAPSR-----RSDYRLVLT 114
Db 59 ELGGERVIVEHARGPRP---DRDGYSGRSGGGYSSRRSTSGTKDYKGPVPTVEYRLIVE 115

QY 115 GLPPSASWQDLKHMKAGDVCFSEVPDRKMGSVVDYSDYDMKYAIRKLDATEP--- 171
Db 116 NLSSRCSWQDLKDFMQAGEVYADAHKRTN-EGVIEFASVDMKRALDLDTGTINGR 174

QY 172 -----RNAFSAYIRVR-----EYESVESRSPDCKSYRSRSGRGP 209
Db 175 NIRLIEDKPTSHRRYSYSGSRSRSRRRSRSSRRSRSSRSRSISKSRS-RRSRSRSGR 233

QY 210 SCYSYS--KSRSVSPARSISPR---SRPLSRSLSYSVSRSGSLARA-----GMIWISQ 258
Db 234 SRSRSKGRSKSKSKKPSKSDRGSHSHSRSKDEYEK-SRGRSRSRSPKENGKDIKSK 292

QY 259 SRSKSRSRSRNSPVSVP 275
Db 293 SRSRSQSRNSPLPVP 309

RESULT 5
SFR4_HUMAN
ID_SFR4_HUMAN STANDARD; PRT; 494 AA.
AC Q08170; Q9BUA4; Q9UEB5;
DT 01-FEB-1995 (Rel. 31, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Splicing factor, arginine/serine-rich 4 (Pre-mRNA splicing factor
DE SRP75) (SRP001LB).
OS SFRS4 OR SRP75.
OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

[1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RP MEDLINE=93309435; PubMed=8321209;

RA Zahler A.M., Neugebauer K.M., Stolk J.A., Roth M.B.;

RT "Human SR proteins and isolation of a cDNA encoding SRp75.";

RL Mol. Cell. Biol. 13:4023-4028 (1993).

[2]
RP SEQUENCE FROM N.A.

RP TISSUE=Lymph;

RP MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Joquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Faley J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Kryzinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.F., Jones S.J.M., Marra M.A.,

RT "Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

[3]
RP SEQUENCE OF 1-192 FROM N.A.

RA Connolly K.S., Gunning K.M., Davis C.A., Kadner K., Subramanian S.,

RA Miguel T., Lewis K.D., Fridlyand J., Alcivar D., Benke J.A.,

RA Boudoc M., Bowen E., Chiang A., Critz P., Jaklevic M.A., Lindo K.,

RA Lindquist K., Miller C., Patel S., Piscia C., Riley B.E., Rojeski H.,

RA Samiento R., Yu C., Montenegro M., Aerts A., Chung A., Abrajano A.,

RA Baker M., Gau C., Jett J., KO C., Beall K., Woolley J.P., Comboy J.,

RA Fang J.F., Narla M., Scultz J.L., Kimmery W., Martin C.H.;

RT "Sequencing of human chromosome 1.";

RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.

[4]
RP SEQUENCE OF 35-45; 84-89; 126-137; 140-154 AND 172-179.

RX MEDLINE=92249775; PubMed=1577277;

RA Zahler A.M., Lane W.S., Stolk J.A., Roth M.B.;

RT "SR proteins: a conserved family of pre-mRNA splicing factors.";

RL Genes Dev. 6:837-847 (1992).

CC -!- FUNCTION: A PROBABLE ROLE IN ALTERNATIVE SPLICE SITE SELECTION

CC -!- DURING PRE-MRNA SPLICING.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES IN THE RS

CC DOMAIN.

CC -!- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.

CC -!- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions. There are no restrictions on way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; L14076; AAA36649.1; --

DR EMBL; BC002781; AA020781.1; --

DR EMBL; AC004236; AAC04476.1; --

DR PIR; A48133; A48133.

DR Genew; HGNC:10786; SFRS4.

DR GK; Q08170; --

DR MIM; 601940; --

DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0008248; F:pre-mRNA splicing factor activity; TAS.
DR GO; GO:0006371; F:mRNA splicing; TAS.

DR InterPro; IPR000504; RNA_rec_mot.

DR Pfam; PF00076; rrm; 2.

DR SMART; SM00360; RRM; 2.

DR PROSITE; PS0102; RRM; 2.

DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.

KW mRNA processing; mRNA splicing; Nuclear protein; RNA-binding;

KW Repeat; Phosphorylation.

FT DOMAIN 2 72

FT DOMAIN 72 97

FT DOMAIN 104 177

FT DOMAIN 179 494

FT CONFLICT 35 35

FT CONFLICT 253 253

FT CONFLICT 318 322

FT CONFLICT 338 338

FT CONFLICT 356 356

FT CONFLICT 436 438

FT CONFLICT 494 AA; 56678 MW; SBBAB917C218C20A CRC64;

SQ SEQUENCE 494 AA; 56678 MW; SBBAB917C218C20A CRC64;

Query Match 29.7%; Score 431; DB 1; Length 494;

Best Local Similarity 40.1%; Pred. No. 1.9e-26;

Matches 124; Conservative 39; Mismatches 92; Indels 54; Gaps 10;

QY 9 IVVGNLPGDIRKCEVEDLIFYKYGIVDIDLKIPRPPGYAFVEFDPDADDAIYGRDGY 68

DB 4 VYIGRLSYCAERDVERFFKYGKILEVDLK-----NGYGFVEFDLFDADDAVYELNGK 58

QY 69 FDFGRLVEIHAHGRSPSPVDYSSYSASRA-----PSRSSDYELVLTGLPPPSAW 122

DB 59 DLGGERVIVEHARGPR-DGSYSGRSGYRRSGDKYGPPTTXYELIVENLSRRCSW 117

QY 123 QDLKDMRKAGDVCFSEVPDRKGMGVYDYSNDMMKYAIRKLDATFR----- 172

DB 118 QDLKDMRKAGDVYADAHKGRKI-EGVIEFVSVDMMKALEKLDGTEWGRKRLVEDK 176

QY 173 --NAFSSAIRVREYESSRSVS-----PDDSKSYRSRSRSGSPSCSYS-SKSRSV 220

DB 177 PGRRRRSYSRSH-SRSRSRHSRKSRSRSGSKSHSKRSRSRSGSRSKSRSR 235

QY 221 SPARSISPR-----SRPLSRSLYSVSRSGSLLRAGDWI-----SQSRSK 262

DB 236 SQSRSKSKSRSPSKSRSHSAGKSRSGKQDAEEXIQNNNVNPKPSRSPSRHK 295

QY 263 SPSRSRSNS 271

DB 296 SKSRSKSRKS 304

RESULT 6

SFR4_MOUSE

ID SFR4_MOUSE STANDARD; PRT; 489 AA.

AC Q8VB97; Q9JJC3;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE SPLICING factor, arginine/serine-rich 4.

GN SFRS4.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.J., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.A., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

SEQUENCE OF 147-489 FROM N.A.

STRAN=57BL/6;

Osada N., Kusuda J., Tanuma R., Ito A., Hirata M., Sugano S., Hashimoto K.;

"Isolation of full-length cDNA clones from mouse brain cDNA library made by oligo-capping method.";

Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.

-!- FUNCTION: A PROBABLE ROLE IN ALTERNATIVE SPLICING

-!- DURING PRE-MRNA SPLICING (By similarity).

-!- SUBCELLULAR LOCATION: Nuclear (By similarity).

-!- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES IN THE RS DOMAIN (By similarity).

-!- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.

-!- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

EMBL; BC019437; AAH19437.1; -

EMBL; AB041587; BAA95070.1; ALT_INIT.

MGI; MGI:1890577; Sfrs4.

InterPro; IPR000504; RNA_rec_mot.

Pfam; PF00076; rrm; 2.

SMART; SM00360; RRM; 2.

PROSITE; PS00102; RRM; 2.

PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.

mRNA processing; mRNA splicing; Nuclear protein; RNA-binding; Repeat; Phosphorylation.

FT DOMAIN 2 72 RNA-BINDING (RRM) 1.

FT DOMAIN 72 97 GLY-RICH (HINGE REGION).

FT DOMAIN 104 177 RNA-BINDING (RRM) 2.

FT DOMAIN 180 489 ARG/SER-RICH (RS DOMAIN).

FT CONFLICT 341 341 S -> SKVGS (IN REF. 2).

FT CONFLICT 389 390 MISSING (IN REF. 2).

FT CONFLICT 421 421 G -> E (IN REF. 2).

SEQUENCE 489 AA; 55979 MW; 8D5FE8D1EF4624B3 CRC64;

Query Match

Best Local Similarity 33.1%; Score 426.5; DB 1; Length 489;

Matches 120; Conservative 44; Mismatches 78; Indels 121; Gaps 10;

QY 9 IYVGNLPDTRKCEVEDLFPKYGPVIDIDLKIPRPGYAFVEFEDPRDADDAIYGRDGY 68

Db 4 VYIGRLSYQARERDVEFFFGYKILVDLK-----NGYGFVEFDDURDADDAIYELNGK 58

QY 69 DFDGRLRVIAHGGRFSPSDRYSSSYASRA-----PSRRSDYRVLVATG 115

Db 59 DLGSRVIVHARGPRR-----DSYSGSRSGYGRSGRDKYGFPTTEYRLIVEN 110

QY 116 LPFSASQDLDKHNKAGDYCFSEVFPDRKMGSGVDYSYDDMKYAIRKLDATF---- 171

Db 111 LSSRCSNQDLKYVVRQAGEVTYADAHGRKN-EGVIEFVSYSMDKRALEKLDGTENVGRK 169

QY	172	-----RNAPS-----	SAIVRYE-----	YSR 180
		:	:	
DB	170	IRLVEDPGRRRRSYGRSHSRGRSHSRKSRSGSKSHSKSRSGSHSR		229
QY	199	SVSVPDDSKYVRGRSRSGPSCYSKSRVSP		225
DB	230	SKSRSRSGSRSKSRKSRSPKDNKSRGRSRSPDKSRKSKDHAEDKLQNDSDAGKKS		289
QY	226	ISP	-----RSRPLSRSRYSVSRSGSLLRAGDWISQSRKSRSR	266
DB	290	HSPRHDSKSRSGOERRAEERRSVSRARSGEKSQEKSLK	-----SRSRSRSRSR	344
QY	267	SRS 269		
DB	345	SRS 347		

RESULT 7

ID	SR55_DROME	STANDARD;	PRT;	375 AA.
DT	P26686;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Serine-arginine protein 55 (SRP55) (Enhancer of deformed) (52-kDa bracketing protein) (B52 protein).			
DE	B (DFP) OR SR55 OR B52.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]	SEQUENCE FROM N.A. AND SEQUENCE OF 1-14; 125-131 AND 136-147.		
RC	STRAIN=CL; TISSUE=Embryo;			
RC	MEDLINE=92011900; PubMed=1717489;			
RA	Roth M.B., Zahler A.M., Stolk J.A.;			
RT	"A conserved family of nuclear phosphoproteins localized to sites of polymerase II transcription."			
RL	J. Cell Biol. 115:587-596(1991).			
RN	[2]	SEQUENCE FROM N.A. (ISOFORM LONG).		
RC	TISSUE=Embryo;			
RC	MEDLINE=91357476; PubMed=1885003;			
RA	Chaplin D.T., Frasch M., Saumweber H., Lis J.T.;			
RT	"Characterization of a Drosophila protein associated with boundaries of transcriptionally active chromatin."			
RL	Genes Dev. 5:1611-1621(1991).			
RN	[3]	CHARACTERIZATION.		
RP	MEDLINE=92159037; PubMed=1741384;			
RA	Mayeda A., Zahler A.M., Krainer A.R., Roth M.B.;			
RT	"Two members of a conserved family of nuclear phosphoproteins are involved in pre-mRNA splicing."			
RL	Proc. Natl. Acad. Sci. U.S.A. 89:1301-1304(1992).			
RN	[4]	FUNCTION.		
RP	MEDLINE=95021280; PubMed=7935465;			
RA	Ring H.Z., Lis J.T.;			
RT	"The SR protein B52/SRP55 is essential for Drosophila development."			
RL	Mol. Cell. Biol. 14:7499-7506(1994).			
CC	-!	FUNCTION: ESSENTIAL FOR DEVELOPMENT. MAY HAVE A CRITICAL ROLE IN SPLICING OR IN CONTROLLING ALTERNATIVE SPLICING SITE USE OF AT LEAST SOME PRE-mRNA IN VIVO. NOT REQUIRED FOR ALL SPLICING. MAY PLAY A GENERAL ROLE IN THE CONDENSATION OR DECONDENSATION OF CHROMATIN.		
CC	-!	SUBCELLULAR LOCATION: NUCLEAR; ASSOCIATED WITH BOUNDARIES OF TRANSCRIPTIONALLY ACTIVE CHROMATIN.		
CC	-!	ALTERNATIVE PRODUCTS:		
CC	Event=Alternative splicing; Named isoforms=2;			
CC	Name=Long;			
CC	IsCID=P26686-1; Sequence=Displayed;			
CC	Name=Short;			
CC	IsCID=P26686-2; Sequence=VSP 005878;			


```

RESULT 9
SPR5_RAT
ID      SPR5_RAT      STANDARD;      PRT;      269 AA.
ID      AC      Q09167;  Q35335;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Splicing factor, arginine/serine-rich 5 (Pre-mRNA splicing factor
DE      SRP40) (Insulin-induced growth response protein CL-4) (Delayed-early
DE      protein HRS).
GN      SPR55 OR CL-4 OR HRS.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC      NCBI_TaxID=10116;
[1]
RN      SEQUENCE FROM N.A.
RP      TISSUE=Liver;
RC      MEDLINE=93315501; PubMed=7686911;
RX      Diamond R.H., Du K., Lee V.M., Mohn K.L., Haber B.A., Tewari D.S.,
RA      Taub R.;
RT      "Novel delayed-early and highly insulin-induced growth response
RT      genes. Identification of HRS, a potential regulator of alternative
RT      pre-mRNA splicing.";
RL      J. Biol. Chem. 268:15185-15192(1993).
[2]
RN      SEQUENCE FROM N.A. (ISOFORM 2).
RP      TISSUE=Thymus;
RC      MEDLINE=98094279; PubMed=9434190;
RX      Du K., Taub R.;
RA      "Alternative splicing and structure of the human and mouse
RT      SPR55/HRS/Srp40 genes.";
RL      Gene 204:243-249(1997).
[3]
RN      SEQUENCE OF 185-269 FROM N.A.
RP      STRAIN=Sprague-Dawley; TISSUE=Testis;
RC      MEDLINE=94164020; PubMed=8161377;
RX      Hamil K.G., Hall S.H.;
RA      "Cloning of rat Sertoli cell follicle-stimulating hormone primary
RT      response complementary deoxyribonucleic acid: regulation of TSC-22
RT      gene expression.";
RL      Endocrinology 134:1205-1212(1994).
-1- FUNCTION: MAY BE REQUIRED FOR PROGRESSION THROUGH G1 AND ENTRY
CC      INTO S PHASE OF CELL GROWTH. MAY PLAY A REGULATORY ROLE IN PRE-
CC      MRNA SPLICING. AUTOREGULATES ITS OWN EXPRESSION. PLAYS A ROLE IN
CC      CONSTITUTIVE SPLICING AND CAN MODULATE THE SELECTION OF
CC      ALTERNATIVE SPICE SITES (BY SIMILARITY).
CC      -1- FUNCTION: COULD PLAY AN IMPORTANT ROLE IN DEVELOPMENT AND
CC      DIFFERENTIATION IN THE SPLEEN AND THYMUS.
CC      -1- ALTERNATIVE PRODUCTS:
CC      Event=alternative splicing; Named isoforms=2;
CC      Name=1;
CC      IsoId=Q09167-1; Sequence=Displayed;
CC      Name=2; Synonyms=HRR-LF;
CC      IsoId=Q09167-2; Sequence=VSP_005867, VSP_005868;
CC      -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SPLEEN AND THYMUS.
CC      -1- INDUCTION: BY INSULIN AND HEPATECTOMY.
CC      -1- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES IN THE RS
CC      DOMAIN (BY SIMILARITY).
CC      -1- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.
CC      -1- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC      or send an email to license@isb-sib.ch).
-----
DR      ENBL; L13635; AAA62266.1; -
DR      ENBL; AF020683; AAB71864.1; -
DR      ENBL; L33267; AAA42316.1; -

```



```

[2]
SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
TISSUE=Brain, Cervix, Prostate, and Skin;
MEDLINE=22398257; PubMed=12477932;
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustin T.B., Iosifovski S., Carninci P., Frange C.,
Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
FAhey J., Helton E., Kettman M., Madañ A., Rodríguez S., Sánchez A.,
Whitting M., Madañ A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodríguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
Schnerch A., Schein J.F., Jones S.J.M., Marra M.A.;
"generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
-!- FUNCTION: REQUIRED FOR PRE-MNRA SPLICING. CAN ALSO MODULATE
ALTERNATIVE SPLICING IN VITRO.
-!- SUBCELLULAR LOCATION: Nuclear.
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=3;
Comment=Isoforms, often lacking the RS domain and differentially
expressed in fetal tissues, may be involved in modulation of 988
function;
Name=1;
IsoId=Q16629-1; Sequence=Displayed;
Name=2;
IsoId=Q16629-2; Sequence=VSP_005872, VSP_005873;
Name=3;
IsoId=Q16629-3; Sequence=VSP_005874, VSP_005875;
-!- TISSUE SPECIFICITY: BRAIN, LIVER, KIDNEY AND LUNG.
-!- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES IN THE RS
DOMAIN.
-!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
-!- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.
-!- SIMILARITY: Contains 1 CCHC-type zinc finger.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announcements
or send an email to license@isb-sib.ch).
-----
EMBL; L22253; AAA35495.1; --
EMBL; L41887; AAA88098.1; --
DR DR EMBL; BC000997; AAA00997.1; --
DR DR EMBL; BC017369; AAH17369.1; --
DR DR EMBL; BC017908; AAH17908.1; --
DR DR EMBL; BC022328; AAH22328.1; --
DR DR PIR; A57198; A57198.
DR DR HSPG; P11940; LCYJ.
DR DR Genew; HGNC:10789; SFR97.
DR DR GK; Q16629; --
DR DR GO; GO:0005634; C:nucleus; TAS.
DR DR

```

[1]
PP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=9608180; PubMed=9421507;
RA Gross T., Richter K., Mierke C., Luetzelberger M., Kaeufer N.F.;
RT "Identification and characterization of spl1, a gene of fission yeast
RT encoding a RNA binding domain and a RS domain typical of SR splicing
RT factors.";
RL Nucleic Acids Res. 26:505-511(1998).
[2]
PP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21843401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins S., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Money P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voickart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleau V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Foreburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Snpakovski G.V., Usery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
[3]
PP SEQUENCE OF 5-365 FROM N.A.
RC STRAIN=PR745;
RX MEDLINE=98162722; PubMed=9501991;
RA Yoshioka S., Kato K., Nakai K., Okayama H., Nojima H.;
RT "Identification of open reading frames in Schizosaccharomyces pombe
RT cDNAs.";
RL DNA Res. 4:363-369(1997).
[4]
PP FUNCTION: Has a role in pre-mRNA splicing where it is involved in
PP spliceosome assembly (By similarity).
[5]
PP SUBCELLULAR LOCATION: Nuclear (Potential).
[6]
PP PTM: Extensively phosphorylated on serine residues in the RS
PP domain (By similarity?).
[7]
PP SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.
[8]
PP SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.
[9]
PP This SWISS-PROT entry is copyright. It is produced through a collaboration
PP between the Swiss Institute of Bioinformatics and the EMBL outstation -
PP the European Bioinformatics Institute. There are no restrictions on its
PP use by non-profit institutions as long as its content is in no way
PP modified and this statement is not removed. Usage by and for commercial
PP entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
PP or send an email to license@isb-sib.ch).
[10]
PP EMBL; AF012278; CAB39357.1; -.
PP EMBL; AL121745; CAB57400.1; -.
PP EMBL; D89163; BAA13825.1; -.
PP PIR; T37730; T37730.
PP PIR; T42525; T42525.
PP GeneDB SPombe; SPAC16.02c; -.
PP InterPro; IPR000504; RNA_rec_mot.
PP Pfam; PF00076; rrm; 2.
PP SMART; SM00360; RRM; 2.
PP PROSITE; PS50102; RRM; 2.

DR PROSITE; PS00030; RRM RNP 1; FALSE NEG.
KW mRNA processing; mRNA_splicing; Nuclear protein; RNA-binding;
KW Repeat; Phosphorylation.
FT DOMAIN 6 69 RNA-BINDING (RRM) 1.
FT DOMAIN 74 91 GLY-RICH (HINGE REGION).
FT DOMAIN 100 166 RNA-BINDING (RRM) 2.
FT DOMAIN 182 300 ARG/SER-RICH (RS DOMAIN).
FT SEQUENCE 365 AA; 42566 MW; 67E988573A736691 CRC64;
Query Match 19.3%; Score 280; DB 1; Length 365;
Best Local Similarity 31.6%; Pred. No. 8e-15;
Matches 85; Conservative 38; Mismatches 114; Indels 32; Gaps 7;
QY 9 IYVGNLPDGRKCEVEDLFYKYGIVDIDILKIPRPPCYAVEREDPRDADAIYGRDGY 68
DB 6 LFGVRIIPQATREDMDFFKYGQILCKLM-----NGFGEVEVEDADADIYVDFGK 60
QY 69 DFDGRLRVEIAHGRRRSPVDYSSYSASRAPSRRSDYRVLVTGLPPSASWQDLKDH 128
DB 61 EFMGSRIVVEPARGERR---RRENFRESAASKYPRPRTGRLIVENLSEDSVQDLKDV 117
QY 129 MRKAGDVCFSEVFPDRKMGSGVDYSNYDDMKYAIRKLD-----ATEFNAFSAVIR 181
DB 118 MRKAGEPFTDAHRENPQ-AGVVEFSTEEDVRNALTSLNGEVIKQAVTLREDPDAAPEP 176
QY 182 VREYRSVRSVSPDDSKYRSRSRSGPSCYSKSRSVSPARSISPRSRPLSRSLYS 241
DB 177 LPEVPSRFRSPSPARRRYRDYR-RGGDYRRDAYRFGDRDDERYAPRGEYRNRDEY- 234
QY 242 SVSRSGSLLRAGDMISQSRKSRSRSRN 270
DB 235 -----RRG-----GRDYRRNRSRSD 249
RESULT 13
SPR2_CHICK STANDARD; PRT; 221 AA.
AC P30352;
DT 01-APR-1993 (Rel. 25; Created)
DT 01-APR-1993 (Rel. 25; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Splicing factor, arginine/serine-rich 2 (Splicing factor SC35) (SC-35)
DE (Splicing component, 35 kDa) (PR264 protein).
GN SFRS2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus.
RX MEDLINE=92212859; PubMed=1557353;
RA Vellard M., Sureau A., Soret J., Martinerie C., Perbal B.;
RT "A potential splicing factor is encoded by the opposite strand of the
RT trans-spliced c-myc exon.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2511-2515(1992).
CC -!- FUNCTION: NECESSARY FOR THE SPLICING OF PRE-MENA. IT IS REQUIRED
CC FOR FORMATION OF THE EARLIEST ATP-DEPENDENT SPLICING COMPLEX AND
CC INTERACTS WITH SPLICEOSOMAL COMPONENTS BOUND TO BOTH THE 5' AND 3'
CC SPICE SITES DURING SPLICEOSOME ASSEMBLY. IT ALSO IS REQUIRED FOR
CC ATP-DEPENDENT INTERACTIONS OF BOTH U1 AND U2 SNRNPS WITH PRE-
CC MENA.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES IN THE RS
CC DOMAIN (BY SIMILARITY).
CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -!- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).

EMBL; X62446; CAA44306.1; --
PIR; B42701; B42701.
HSSP; P19339; ISXL.
InterPro; IPR000504; RNA_rec_mct.
Pfam; PF00076; rrm; 1.
SMART; SMC0360; RRM; 1.
PROSITE; PS50102; RRM; 1.
PROSITE; PS00030; RRM_RNP_1; 1.
KW Nuclear protein; RNA-binding; mRNA splicing; Phosphorylation.
FT DOMAIN 14 92 RNA-BINDING (RRM).
GLY-RICH (HINGE REGION).
FT DOMAIN 111 116 ARG/SER-RICH (RS DOMAIN).
FT DOMAIN 117 221
SQ SEQUENCE 221 AA; 25524 MW; 75A4D8FF9170F1BF CRC64;

Query Match 16.4%; Score 237.5; DB 1; Length 221;
Best Local Similarity 31.5%; Pred. No. 8.9e-12;
Matches 86; Conservative 24; Mismatches 60; Indels 103; Gaps 10;

Qy 8 TIYGVN-----PDIRKCEVEDLFYKIGVIDL---KIIPPPGYAFVEEDPRDA 59
Db SLKYDNLTYSPTDLR----VFKEYGRVGDVIIDRYTKESRGFAVRFDKDEAE 69
Qy 60 DAIXGRGCVDFDGCLRAIEAHGRRFPSPVDRIYSSSYASRAPSRSDRYVLVTGLPPS 119
Db DAWDAMDGAIVLDGEILVQWARYGR-----PPD 97
Qy 120 ASMQDLKDHRKAGDVCFSEVPDRKGMSGVVDSNYDDMKYAIRKLDATEFRNAFSAY 179
Db 98 S-----HHSKRGPPP-----PRYGSG-- 114
Qy 180 IRVREYERSVSRRPSDPDSKSVYRSRRSRGRGPCSVSSKSRSVSVPARSISPRSRLSRSL 239
Db 115 -----YGR--SRSPRRRRSRSSRSRRSRSSRSYRSKRSTRS-RSRSTKSRSA 166
Qy 240 YSSVSRSGSLLRADWTISOGRSKRSRSNSRP 272
Db 167 RRKSXSSS-----VSRGSRRSRRSRSP 192

RESULT 14
SF2_MOUSE STANDARD; PRF: 221 AA.

AC Q62092;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Splicing factor, arginine/serine-rich 2 (Splicing factor SC35) (SC-35)
DE (Splicing component, 35 kDa) (PR264 protein).
GN SPF52 OR PR264.
OS Mus musculus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=38447613; PubMed=9774382;
RA Yang L., Embree L.J., Tsai S., Hickstein D.D.;
RT "Oncoprotein TLS interacts with serine-arginine proteins involved in
RNA splicing."
RL J. Biol. Chem. 273:27761-27764(1998).
RN [2]

SEQUENCE OF 1-121 FROM N.A.
RC STRAIN=129/sv; TISSUE=Liver;
RC Gaillard C., Perbal B.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NECESSARY FOR THE SPLICING OF PRE-MRNA. IT IS REQUIRED
FOR FORMATION OF THE EARLIEST ATP-DEPENDENT SPLICING COMPLEX AND
INTERACTS WITH SPLICEOSOMAL COMPONENTS BOUND TO BOTH THE 5' AND
3' SPLICING SITES DURING SPLICEOSOME ASSEMBLY. IT ALSO IS REQUIRED
FOR THE FORMATION OF A SECONDARY SPLICING SITE.
CC

```

CC FOR ATP-DEPENDENT INTERACTIONS OF BOTH U1 AND U2 SNRNPS WITH PRE-
CC MRNA (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES IN THE RS
CC DOMAIN (BY SIMILARITY).
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -1- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; AR077858; AAC71000.1; -.
DR ENBL; X98511; CAA67134.1; -.
DR HSP; P19339; 1SXL.
DR MGD; MGI:98284; Sfrs2.
DR GO; GO:0005681; C:spliceosome complex; IDA.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS0102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
DR Nuclear protein; RNA-Binding, mRNA splicing; Phosphorylation.
FT DOMAIN 14 92
FT FT 111 116 GYI-RICH (HINGE REGION).
FT FT 117 221 ARG/SER-RICH (RS DOMAIN).
FT SQ SEQUENCE 221 AA; 25476 MW; 68121AC4D35714FA CRC64;
Query Match 16.2%; Score 234.5; DB 1; Length 221;
Best Local Similarity 31.6%; Pred. No. 1.5e-11;
Matches 87; Conservative 24; Mismatches 63; Indels 101; Gaps 10;
Qy 8 TIYVGNL-----PGDIRKCEVDLFKYGPVIDL---KIPRPDGYAFVFEDPRDAD 59
Db 15 SLKVDNLITRSPDILR-----VFEKYRGVDVTPDRYTKESGFAFVPHDKDAE 69
Qy 60 DALYGRDGVDFGCRURVIAHGRRRFPSPVDYSSYSASRAPSRSDYRLVLTGLPPS 119
Db 70 DAMDAMDGAVLDGRELQVMQARYR-----PPD 97
Qy 120 ASWQDLKDHWRAGDVCFSEVPEDKMGSVVDYNYDDMKYAIRKLDATFENAFSSAY 179
Db 98 S-----HHSRGGP-----PRYGGG----- 114
Qy 180 IRVREYESRSRSPDSSKYSRSRSGRSPSCYSKSRSPARSISPRSRPLGSRSL 239
Db 115 -----YGR--SRSPRRRSRSPSRSPSRSPSRSPSRSPSRSPSRSPSRSPSR 166
Qy 240 YGSVRSGLLRAGDWISQSRKSRSPSRSPSPS 274
Db 167 RRSKSSSVSR-----SRSRSPSRSPSRSPSPSPS 196
RESULT 15
SFR3_HUMAN STANDARD; PRT; 164 AA.
ID SFR3_HUMAN
AC P23152; C08831;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Splicing factor, arginine/serine-rich 3 (Pre-mRNA splicing factor
DE SRP20) (X16 protein).
OS SFRP3 OR SRP20 OR X16.
GN Homo sapiens (Human), and
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606, 10090;
RN [1]

```

RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-37 AND 54-64.
RC SPECIES=Human;
RX MEDLINE=92249775; PubMed=1577277;
RA Zahler A.M., Lane W.S., Stolk J.A., Roth M.B.;
RT "SR proteins: a conserved family of pre-mRNA splicing factors.";
RL Genes Dev. 6:837-847(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RA Liu W.L., Wang M., Tang D., Rodgers G.;
RT "Identification and characterization of novel full-length cDNAs
RT differentially expressed in hematopoietic lineages";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalka U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse;
RX MEDLINE=9123908; PubMed=2030943;
RA Ayane M., Preuss U., Koehler G., Nielsen P.J.;
RT "A differentially expressed murine RNA encoding a protein with
RT similarities to two types of nucleic acid binding motifs";
RL Nucleic Acids Res. 19:1273-1278(1991).
RN [5]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC SPECIES=Mouse; STRAIN=MRL; TISSUE=Lymphoid;
RX MEDLINE=97299661; PubMed=9154810;
RA Jumaa H., Guenet J.-L., Nielsen P.J.;
RT "Regulated expression and RNA processing of transcripts from the
RT Sp20 splicing factor gene during the cell cycle";
RL Mol. Cell. Biol. 17:3116-3124(1997).
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse; STRAIN=C57BL/6J;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavanta H.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Anon H., Baldarelli R., Bash G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- FUNCTION: MAY BE INVOLVED IN RNA PROCESSING IN RELATION WITH
CC CELLULAR PROLIFERATION AND/OR MATURATION.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=P23152-1; Sequence=D:played;
CC Name=Short;
CC IsoId=P23152-2; Sequence=VSP_005861, VSP_005862;
CC Note=Has been shown to exist only in mouse so far;
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN THYMUS AND PRE-B CELL
CC LINES; HIGH, IN TESTIS, BRAIN AND SPLEEN; VERY LOW IN HEART AND
CC NOT DETECTABLE IN LIVER AND KIDNEY.
CC -!- INDUCTION: BY SERUM; IN A TISSUE CULTURE.
CC -!- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES IN THE RS
CC DOMAIN (BY SIMILARITY).
CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -!- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; L10838; AAA36648.1; -;
CC EMBL; AF107405; AAD44523.1; -;
CC EMBL; BC000914; AAH00914.1; -;
CC EMBL; X53824; CAA37821.1; -;
CC EMBL; X91556; CAA62844.1; -;
CC EMBL; X91556; CAA62845.1; -;
CC EMBL; AK011657; BAB27762.1; -;
CC PIR; I54089; I54089.
CC PIR; S14016; S14016.
CC HSP; F11340; IGVJ.
CC Genew; HGNC:10785; SFRS3.
CC MIM; 603364; -;
CC MIM; P23152; -;
CC MGD; MGI:98285; Sfrs3.
CC GO; GO:0003723; F:RNA binding activity; TAS.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF00076; rtm; 1.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS0102; RRM; 1.
CC PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
CC Nuclear protein; RNA-binding; mRNA splicing; Alternative splicing;
CC Phosphorylation; Repeat.
CC FT DOMAIN 10 83 RNA-BINDING (RRM).
FT DOMAIN 86 164 ARG/SER-RICH (RS DOMAIN).
FT DOMAIN 119 164 2 X APPROXIMATE REPEATS, BASIC.
FT REPEAT 119 133 B-1.
FT REPEAT 149 164 B-2.
FT VARSPLIC 115 124 SPFRSFRSRS -> VTMSLTLTL (in isoform
FT Short).
FT FTId=VSP_005861.
FT Missing (in isoform Short).
FT /FTId=VSP_005862.
SQ SEQUENCE 164 AA; 19329 MW; 02F0A5EE33FF28A0 CRC64;
Query Match 16.1%; Score 234; DB 1; Length 164;
Best Local Similarity 30.2%; Pred. No. 1.2e-11;
Matches 79; Conservative 24; Mismatches 47; Indels 112; Gaps 9;
QY 9 IYVGNLPGDRCRVEDLFYKYGFIVDILKIPRPPGAVFEFEDPRDADALVGRDGY 68
Db 12 VYVGNLGNNGKNTLELAFAGYGYPLRSV--VWVNPFGFAVEFEDPRDADAVRELDGR 69

QY	69	DFDGCRLRVEIAHGGRFPSPVDYSSYSASRAPSRSDYRVLVTGLPPSASWQDLKDH	128
Db	70	TLGCGRVRLVLSNGEKR-----SRRGP-----PPSGRRPRDDY	104
QY	129	WRKAGDVCFSEVFPDRKMGSGVVDYKNYDDMKYAIRKLDATFERNAPSSAIRVREYER	188
Db	105	RRRSP-----PPRR-----	113
QY	189	SVSRSPDDSKSYRSRSGRGPSCSYSKRSVSPARSISPRSRPLSRSRSLYSSVSRSGS	248
Db	114	---RSP-----RRSFER-----SRSLSRDRR---RERSLRERNHKP-----	147
QY	249	LIRAGDWISQSRKSRSRSRN	270
Db	148	-----SRSFERSRSRSRN	161

Search completed: January 29, 2004, 02:50:43
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 02:55:41 ; Search time 53 Seconds
(without alignments)
1358.428 Million cell updates/sec

Title: US-10-014-927-19

Perfect score: 1451

Sequence: 1 MSSRNRTIYGNLPGDIRK.....RSKSRSRSRNSPVSPTSG 279

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL.23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_protist.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1377.5	94.9	268	10 Q9XFR5	Q9xfr5 arabidopsis
2	1343	92.6	256	10 Q9XFR6	Q9xfr6 arabidopsis
3	1159	79.9	237	10 O80496	O80496 arabidopsis
4	1032.5	71.2	207	10 Q949S8	Q949S8 arabidopsis
5	863	59.5	307	10 Q9SP11	Q9sp11 arabidopsis
6	847	58.4	285	10 Q8L7P1	Q8l7p1 arabidopsis
7	845	58.2	270	10 Q9SP14	Q9sp14 arabidopsis
8	845	58.2	276	10 Q9SP15	Q9sp15 arabidopsis
9	845	58.2	289	10 Q9SP12	Q9sp12 arabidopsis
10	827.5	57.0	300	10 Q8H453	Q8h453 oryza sativ
11	806	55.5	235	10 Q9CA06	Q9ca06 arabidopsis
12	748	51.6	250	10 Q8RZS6	Q8rzs6 oryza sativ
13	745	51.3	261	10 Q9SP13	Q9sp13 arabidopsis
14	649.5	44.8	178	10 Q8GXSO	Q8gxso arabidopsis
15	646.5	44.6	294	10 Q8L290	Q8l290 arabidopsis
16	597.5	41.2	243	10 Q9SCL3	Q9scl3 arabidopsis

17	586.5	40.4	283	13 Q8AVB5	Q8avb5 xenopus lae
18	566.5	39.0	258	5 Q9NEW6	Q9new6 caenorhabdi
19	566.5	39.0	258	5 Q9GQ17	Q9gq17 caenorhabdi
20	547	37.7	255	5 Q9V3W7	Q9v3w7 drosophila
21	545.5	37.6	248	11 Q8C1H9	Q8c1h9 mus musculu
22	525	36.2	222	11 Q9DOB0	Q9dob0 mus musculu
23	497	34.3	201	11 Q8BTV3	Q8btv3 mus musculu
24	458	31.6	190	11 Q9CRN3	Q9crn3 mus musculu
25	433.5	29.9	350	5 Q8M266	Q8m266 drosophila
26	432.5	29.8	339	11 Q921K3	Q921k3 mus musculu
27	432.5	29.8	339	11 Q9DBP1	Q9dbp1 mus musculu
28	431	29.7	329	5 Q9VFT0	Q9vft0 drosophila
29	412.5	28.4	270	5 Q8I3T5	Q8i3t5 plasmodium
30	411	28.3	538	5 Q8IJU0	Q8ijj0 plasmodium
31	407.5	28.1	322	5 Q23796	Q23796 chironomus
32	406.5	28.0	346	5 Q8ING9	Q8ing9 drosophila
33	405.5	27.9	270	11 Q9D8S5	Q9d8s5 mus musculu
34	366	25.2	489	11 Q8K3A8	Q8k3a8 mus musculu
35	315.5	21.7	281	5 Q23120	Q23120 caenorhabdi
36	307	21.2	238	11 Q8R3E9	Q8r3e9 mus musculu
37	304.5	21.0	267	11 Q8BLJ97	Q8blj97 mus musculu
38	303.5	20.9	235	11 Q8BUR2	Q8bur2 mus musculu
39	294	20.3	312	5 Q23121	Q23121 caenorhabdi
40	288	19.8	226	11 Q91VS1	Q91vs1 mus musculu
41	258.5	17.8	284	10 Q9FYB7	Q9fyb7 arabidopsis
42	256	17.6	290	10 Q8VTA5	Q8vta5 arabidopsis
43	253	17.4	272	4 Q96TA3	Q96ta3 homo sapien
44	253	17.4	282	4 Q9BRL6	Q9brl6 homo sapien
45	249.5	17.2	262	11 Q9R0U0	Q9r0u0 mus musculu

ALIGNMENTS

RESULT 1

Q9XFR5	PRELIMINARY;	PRT;	268 AA.
ID	Q9XFR5		
AC	Q9XFR5;		
DT	01-NOV-1999 (TRENBLrel. 12, Created)		
DT	01-NOV-1999 (TRENBLrel. 12, Last sequence update)		
DT	01-MAR-2003 (TRENBLrel. 23, Last annotation update)		
DE	SP2/ASF-like splicing modulator Srp30 (Putative SF2/ASF splicing modulator Srp30).		
DE	modulator Srp30).		
GN	Srp30 OR AT1G09140.		
OS	Arabidopsis thaliana (Mouse-ear cross).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99234087; PubMed=10215626;		
RA	Lopato S., Kalyana M., Dörner S., Kobayashi R., Krainer A.R., Barta A.;		
RT	"atsRp30, one of two SF2/ASF-like proteins from Arabidopsis thaliana,		
RT	regulates splicing of specific plant genes.";		
RL	Genes Dev. 13:987-1001(1999).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,		
RA	Quach H.L., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,		
RA	Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,		
RA	Palm C.J., Shinn P., Southwick A., Tripp M.G., Wu T., Davis R.W.,		
RA	Ecker J.R., Theologis A.;		
RT	"Arabidopsis Open Reading Frame (ORF) Clones.";		
RT	Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AJ131214; CAB42557.1; -		
DR	EMBL; AY150486; AAN13011.1; -		
DR	HSSP; P08579; 1A9N.		
DR	InterPro; IPR000504; RNA_rec_mot.		
DR	Pfam; PF00076; rim; 2.		
DR	SMART; SM00360; RRM; 2.		
DR	PROSITE; PS50102; RRM; 2.		
SQ	SEQUENCE 268 AA, 30385 MW; 73BDC3534A8F9AC4 CRC64;		

Query Match		94.9%; Score 1377.5; DB 10; Length 268;
Best Local Similarity		96.1%; Pred. No. 1.4e-114;
Matches 268; Conservative		0; Mismatches 0; Indels 11; Gaps 1;
Qy	1	MSSRWNRITVYGNLPGDIRKCEVEDLFYKGPVIDIDLKIPRPPGYAFVEFEDPRDADD 60
Db	1	MSSRWNRITVYGNLPGDIRKCEVEDLFYKGPVIDIDLKIPRPPGYAFVEFEDPRDADD 60
Qy	61	AIYGRDGYDFDGCRLRVEIAHGGRRFSPVDYSSYSASRAPRRSDYRVLTGLPPSA 120
Db	61	AIYGRDGYDFDGCRLRVEIAHGGRRFSPVDYSSYSASRAPRRSDYRVLTGLPPSA 120
Qy	121	SWQDLKDHMRKAGDVCFSEVPFDRKMGSGVVDYSDYSSYSASRAPRRSDYRVLTGLPPSA 180
Db	121	SWQDLKDHMRKAGDVCFSEVPFDRKMGSGVVDYSDYSSYSASRAPRRSDYRVLTGLPPSA 180
Qy	181	RVREYESRSVSRSPDDSKSYRSRSGPSCSYSKSRSPARSISPRSRPLSRSRSLY 240
Db	181	RVREYESRSVSRSPDDSKSYRSRSGPSCSYSKSRSPARSISPRSRPLSRSRSLY 240
Qy	241	SSVSRSGSLLRAGDWISQSRSKSRSRSGPSPVSVISG 279
Db	241	SSVSR-----SQSRSKSRSRSGPSPVSVISG 268
RESULT 2		
Q9XFR6		
ID	Q9XFR6	PRELIMINARY; PRT; 256 AA.
AC	Q9XFR6	
DT	01-NOV-1999	(TrEMBLrel. 12, Created)
DT	01-NOV-1999	(TrEMBLrel. 12, Last sequence update)
DT	01-OCT-2002	(TrEMBLrel. 22, Last annotation update)
DE	SP2/ASF-like splicing modulator Srp30, variant 1.	
GN	Srp30.	
OS	Arabidopsis thaliana (Mouse-ear cress).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;	
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.	
OX	NCBI_TaxID=3702;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=99234087; PubMed=10215626;	
RA	Lopato S., Kalyna M., Dornier S., Kobayashi R., Krainer A.R., Barta A.;	
RT	"at-Srp30, one of two SP2/ASF-like proteins from Arabidopsis thaliana,	
RT	regulates splicing of specific plant genes."	
RL	Genes Dev. 13:987-1001(1999).	
DR	EMBL; AJ131214; CAB42558.1; -.	
DR	HSSP; P08579; 1A9N.	
DR	InterPro; IPR000504; RNA_rec_mot.	
DR	Pfam; PF00076; rtm; 2.	
DR	SMART; SM00360; RRM; 2.	
DR	PROSITE; PS0102; RRM; 2.	
SQ	SEQUENCE 256 AA; 29113 MW; 834FE60E6E8FC25E CRC64;	
Query Match		92.6%; Score 1343; DB 10; Length 256;
Best Local Similarity		100.0%; Pred. No. 1.6e-111;
Matches 256; Conservative		0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	MSSRWNRITVYGNLPGDIRKCEVEDLFYKGPVIDIDLKIPRPPGYAFVEFEDPRDADD 60
Db	1	MSSRWNRITVYGNLPGDIRKCEVEDLFYKGPVIDIDLKIPRPPGYAFVEFEDPRDADD 60
Qy	61	AIYGRDGYDFDGCRLRVEIAHGGRRFSPVDYSSYSASRAPRRSDYRVLTGLPPSA 120
Db	61	AIYGRDGYDFDGCRLRVEIAHGGRRFSPVDYSSYSASRAPRRSDYRVLTGLPPSA 120
Qy	121	SWQDLKDHMRKAGDVCFSEVPFDRKMGSGVVDYSDYSSYSASRAPRRSDYRVLTGLPPSA 180
Db	121	SWQDLKDHMRKAGDVCFSEVPFDRKMGSGVVDYSDYSSYSASRAPRRSDYRVLTGLPPSA 180
Qy	181	RVREYESRSVSRSPDDSKSYRSRSGPSCSYSKSRSPARSISPRSRPLSRSRSLY 240
Db	181	RVREYESRSVSRSPDDSKSYRSRSGPSCSYSKSRSPARSISPRSRPLSRSRSLY 240

Db	181	RVREYESRSVSRSPDDSKSYRSRSGPSCSYSKSRSPARSISPRSRPLSRSRSLY 240
Qy	241	SSVSRSGSLLRAGDWI 256
Db	241	SSVSRSGSLLRAGDWI 256
RESULT 3		
O80496		
ID	O80496	PRELIMINARY; PRT; 237 AA.
AC	O80496	
DT	01-NOV-1998	(TrEMBLrel. 08, Created)
DT	01-NOV-1998	(TrEMBLrel. 08, Last sequence update)
DT	01-OCT-2002	(TrEMBLrel. 22, Last annotation update)
DE	Tl2M4.19 protein.	
GN	Tl2M4.19	
OS	Arabidopsis thaliana (Mouse-ear cress).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;	
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.	
OX	NCBI_TaxID=3702;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ev. Columbia;	
RA	Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Oji, O, Kwan, A.,	
RA	Liu S., Li J., Araujo R., Au M., Brendel V., Buehler E., Conway A.B.,	
RA	Conway A.R., Dewar K., Feng K., Kim C., Kurtz D., Li Y., Palm C.J.,	
RA	Shinn P., Sun H., Davis R.W., Ecker J.R., Federspiel N.A.,	
RA	Theologis A.;	
RT	"Arabidopsis thaliana chromosome 1 BAC Tl2M4 sequence, complete	
RT	sequence."	
RL	Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ev. Columbia;	
RA	Theologis A.;	
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AC003114; AAC24092.1; -.	
DR	HSSP; P08579; 1A9N.	
DR	InterPro; IPR000504; RNA_rec_mot.	
DR	Pfam; PF00076; rtm; 2.	
DR	SMART; SM00360; RRM; 2.	
DR	PROSITE; PS0102; RRM; 2.	
SQ	SEQUENCE 237 AA; 27438 MW; AF63E0FF1E274F9E CRC64;	
Query Match		79.9%; Score 1159; DB 10; Length 237;
Best Local Similarity		100.0%; Pred. No. 3.4e-95;
Matches 218; Conservative		0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	MSSRWNRITVYGNLPGDIRKCEVEDLFYKGPVIDIDLKIPRPPGYAFVEFEDPRDADD 60
Db	1	MSSRWNRITVYGNLPGDIRKCEVEDLFYKGPVIDIDLKIPRPPGYAFVEFEDPRDADD 60
Qy	61	AIYGRDGYDFDGCRLRVEIAHGGRRFSPVDYSSYSASRAPRRSDYRVLTGLPPSA 120
Db	61	AIYGRDGYDFDGCRLRVEIAHGGRRFSPVDYSSYSASRAPRRSDYRVLTGLPPSA 120
Qy	121	SWQDLKDHMRKAGDVCFSEVPFDRKMGSGVVDYSDYSSYSASRAPRRSDYRVLTGLPPSA 180
Db	121	SWQDLKDHMRKAGDVCFSEVPFDRKMGSGVVDYSDYSSYSASRAPRRSDYRVLTGLPPSA 180
Qy	181	RVREYESRSVSRSPDDSKSYRSRSGPSCSYSKSR 218
Db	181	RVREYESRSVSRSPDDSKSYRSRSGPSCSYSKSR 218
RESULT 4		
Q949S8		
ID	Q949S8	PRELIMINARY; PRT; 207 AA.
AC	Q949S8	
DT	01-DEC-2001	(TrEMBLrel. 19, Created)
DT	01-OCT-2002	(TrEMBLrel. 22, Last sequence update)
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)

```

DE Putative SF2/ASF splicing modulator Srp30 protein (Fragment).
GN AT1G09140.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Arabidopsis Full Length cDNA Clones.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY050912; AAK93589.2; -.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
FT NON TER
SQ SEQUENCE 207 AA; 23336 MW; 76B74CEC8FD4772 CRC64;

Query Match 71.2%; Score 1032.5; DB 10; Length 207;
Best Local Similarity 94.0%; Pred. No. 5.3e-84;
Matches 205; Conservative 1; Mismatches 1; Indels 11; Gaps 1;

QY 52 IYGRDGYDFDGLRLVEIAHGRFSPSVDRYSYSSASRAPSRSDYRLVTGLPPSAS 121
Db 1 IYGRDGYDFDGLRLVEIAHGRFSPSVDRYSYSSASRAPSRSDYRLVTGLPPSAS 60
QY 122 WQDLKDHMRKAGDVCFSEVFPDRKGMGVDPVSYNDMMKYAIRKLDATFENAFSSA 181
Db 61 WQDLKDHMRKAGDVCFSEVFPDRKGMGVDPVSYNDMMKYAIRKLDATFENAFSSA 120
QY 182 VREYESRSVSPDCKSKYRSRSGRSPSCSYSSKRSVSPARISPRSRPLSRSLYS 241
Db 121 VREYESRSVSPDCKSKYRSRSGRSPSCSYSSKRSVSPARISPRSRPLSRSLYS 180
QY 242 SVSRSGSLLRAGDWISQSRSLRSRSGRSPSPVSPVSG 279
Db 181 SVSR-----SQSRSKSRTRSGRSPSPVSPVSG 207

RESULT 5
Q9SP11 PRELIMINARY; PRT; 307 AA.
AC Q9SP11;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Splicing factor SRI.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Landsberg;
RA Lazar G., Goodman H.M.;
RT "The Arabidopsis splicing factor SRI is regulated by alternative
RT splicing.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF173640; AAD52609.1; -.
DR HSP39; P19339; ISXL.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50102; RRM; 2.

```

```

SQ SEQUENCE 307 AA; 34114 MW; CB51862850967EB6 CRC64;

Query Match 59.5%; Score 863; DB 10; Length 307;
Best Local Similarity 62.2%; Pred. No. 1.1e-88;
Matches 184; Conservative 33; Mismatches 45; Indels 34; Gaps 7;

QY 1 MSSRWNRITVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPRPPGYAFVEFDDPDAAD 60
Db 1 MSSRSRTVTVGNLPGDIREREVEDLFYKYGPIVDIDLKVPFPFGYAFVEFDDAADA 60
QY 61 AIYGRDGYDFDGLRLVEIAHGRFSPSVDRYSYSSASRAPSRSDYRLVTGLPPSAS 103
Db 61 AIHGRDGYDFDGLRLVAVELAHGRR---SSDTRGSGFNGGGRGGRGDRGDSGRGPF 117
QY 104 SRSDYRLVTGLPPSASQDLKDHMRKAGDVCFSEVFPDRKGMGVDPVSYNDMMKYAIR 163
Db 118 SRSEFRVLVTGLPSSASQDLKDHMRKAGDVCFSEVFPDRKGMGVDPVSYNDMMKYAIR 177
QY 164 RKLDATEFENAFSSAIVREYESRSVSPDCKSKY-RSRSRSGPSCSYSSKRSVSP 222
Db 178 KLUDDTEFENAFSGYRVREYDRKDSRSPGRSGRSRSGRSV---SRSRSGR 234
QY 223 ARSISPRSPGLSRSLYSVSRSGSLLRAGDWISQSRSLRSRSGRSPSPVSPV 276
Db 235 SRSRSPKAK--SRSRSPAKSTSRSPGPR-----SKSRSPSRSGRSGRSLPSV 282

RESULT 6
Q8L7P1 PRELIMINARY; PRT; 285 AA.
AC Q8L7P1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE SF2/ASF-like splicing modulator Srp30, putative.
GN AT1G02840.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV128338; AAM91541.1; -.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 285 AA; 31967 MW; 85FC1FCAD24C9BF6 CRC64;

Query Match 58.4%; Score 847; DB 10; Length 285;
Best Local Similarity 61.9%; Pred. No. 2.7e-67;
Matches 177; Conservative 32; Mismatches 41; Indels 36; Gaps 6;

QY 1 MSSRWNRITVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPRPPGYAFVEFDDPDAAD 60
Db 1 MSSRSRTVTVGNLPGDIREREVEDLFYKYGPIVDIDLKVPFPFGYAFVEFDDAADA 60
QY 61 AIYGRDGYDFDGLRLVEIAHGRFSPSVDRYSYSSASRAPSRSDYRLVTGLPPSAS 107
Db 61 AIHGRDGYDFDGLRLVAVELAHGRR---SSDTRGSGFNGGGRGGRGDRGDSGRGPF 117
QY 108 DYRVLVTGLPPSASQDLKDHMRKAGDVCFSEVFPDRKGMGVDPVSYNDMMKYAIRKLD 167
Db 118 EFRVLVTGLPSSASQDLKDHMRKAGDVCFSEVFPDRKGMGVDPVSYNDMMKYAIRKLD 177

```



```
QY 1 MSRRWRTTVVGNLPGDIRKCEVDLFYKYGPVVDIDLKIPRPPGYAFVFEDPRDADD 60
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MSRRSRTTVVGNLPGDIREREVEDLFYKIPVQVVDLKVPRPPGYAFVFEDDARDAD 60
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 AIYGRDGYDFDGCRLRVEIAHGRFRPSPVSDRYSSYS-----ASRAP 103
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 AIHGRDGYDFDGCRLRVELAHGRR---SSDTRGSFNGGGRGGGRDGRDGGSRGP 117
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 104 SRRSDYRVLTGLPSSASWQDLKDHWRKAGVCFSEVPDRKMGSGVVDYNDYDMKYAI 163
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 118 SRRSEFRVLTGLPSSASWQDLKDHWRKAGVCFSEVPDRKMGSGVVDYNDYDMKYAL 177
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 164 RKLDATFEPNASSAYIRVREYSRVSPPDSKSY-RSRSRSGPSCSVSSKRSYSP 222
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 178 KLLDTEFNANSGYRVREYDRKDSRSPGRGSRYSKRSRSGRV---SRSRSR 234
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 223 ARSIPSRPLSRSLYSVSRSGSLLRAGDWISQSRKSRSRSRNSP 272
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 235 SRSRSPKAK-SRRSPAKSTERSPG-----PRSKSRSPSP 268
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 10
QBH453 PRELIMINARY; PRT; 380 AA.
AC Q8H453;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Putative pre-mRNA splicing factor SF2 (SR1 protein).
GN P0470D12.43.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
  clone: P0470D12.43";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF004300; BAC16007.1; -.
SQ SEQUENCE 380 AA; 42656 MW; C0B591F23B597022 CRC64;
  Query Match 57.0%; Score 827.5; DB 10; Length 380;
  Best Local Similarity 65.0%; Pred. No. 2.2e-55;
  Matches 180; Conservative 23; Mismatches 49; Indels 25; Gaps 6;
QY 1 MSRRWRTTVVGNLPGDIRKCEVDLFYKYGPVVDIDLKIPRPPGYAFVFEDPRDADD 60
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 67 MSRRSRTTVVGNLPGDIREREVEDLFYKIPVQVVDLKVPRPPGYAFVFEDDARDAD 126
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 AIYGRDGYDFDGCRLRVEIAHGRFRPSPVSDRYSSYSASR--APSRSDYRVLTGLPP 118
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 127 AIRGRDGYDFDGCRLRVELAHGRRGNSSS--FNNSGGGGRGGVSRTEYRVLTGLPS 183
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 119 SASWQDLKDHWRKAGVCFSEVPDRKMGSGVVDYNDYDMKYARKLDATFEPNASSA 178
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 184 SASWQDLKDHWRKAGVCFSEVPDRKMGSGVVDYNDYDMKYARKLDDBEFKNAFSKA 243
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 179 YRVRYEY---SRVSRSPPDSKSYRSRSGPSCSVSSKRSYSPARSPLSR 235
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 244 YRVKDYDGRSGYSRSGS-----RSRSGRSGYSRSGSPRSGGKSPKGSRRSASR 296
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 236 SRSLSVSRSGSLLRAGDWISQSRKSRSRSRNSP 272
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 297 SRSR-SRRSRSE-----SKGRSPSRSPARSQSP 323
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 11
Q9CA06 PRELIMINARY; PRT; 295 AA.
AC Q9CA06;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Putative splicing factor, 53460-55514.
GN TIG12.13.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haase B.J., Wu D.,
  Maiti R., Rensing C.M., Koo H., Fujii C.Y., Utterback T.R.,
  Bartsch M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
  "Arabidopsis thaliana chromosome 1 BAC TIG12 genomic sequence.";
  Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC012329; AAG52185.1; -.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rim; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS0102; RRM; 2.
SQ SEQUENCE 295 AA; 33070 MW; B257767C2876B9CB CRC64;
  Query Match 55.5%; Score 806; DB 10; Length 295;
  Best Local Similarity 61.8%; Pred. No. 1.3e-63;
  Matches 183; Conservative 28; Mismatches 53; Indels 32; Gaps 10;
QY 1 MSRRWRTTVVGNLPGDIRKCEVDLFYKYGPVVDIDLKIPRPPGYAFVFEDPRDADD 60
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MSORFSGSIYVGNLPGDIRHEIEDIFYKYGRVVDIELKVPFRPPGYCFVEFHSRDAED 60
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 AIYGRDGYDFDGCRLRVEIAHGRFRPSPVSDRY-----SSSY-----SASRPSRR 106
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 AIKRGDGYNDGCRRLRVELAHGRRGQSSS--DRGGYGGGGSGYGGGGGGGSGARFGVSRH 119
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 107 SDYRVLTGLPSSASWQDLKDHWRKAGVCFSEVPDRKMGSGVVDYNDYDMKYAIRKL 166
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 SEFRVIVRGLPSSASWQDLKDHWRKAGVCFSEVPDRKMGSGVVDYNDYDMKYAIRKL 179
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 167 DATEFRNAPSAIRVREYR-SRSVSRSPDSSKSYRSRSGRSGPSCSVS---SKRSRVS 221
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 DDFEFPNPAWGFIRVKYESSRSGRSPSRSGRSGRSGRSGRSGRSGRSGRSGRSGR 239
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 222 PARSI---SPR---SRPLSRSGSLYSVSRSGSLLRAGDWISQSRKSRSRSRNSP 272
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 PRKOLSKSPRSLSRSGRSGS--PSPDKKSPRA-----MSRSGRSGRSGRSGRSP 288
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 12
QBZS6 PRELIMINARY; PRT; 250 AA.
AC QBZS6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Putative pre-mRNA splicing factor SF2.
GN B1012D10.26.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
  clone: B1012D10.26";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; AP003535; BAB90350.1; -
DR Gramene; Q8R2S6; -
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rim; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 250 AA; 28444 MW; ED40BD015D5FA31A CRC64;

Query Match 51.6%; Score 748; DB 10; Length 250;
Best Local Similarity 57.6%; Pred. No. 1.5e-58;
Matches 167; Conservative 22; Mismatches 41; Indels 60; Gaps 7;

Qy 1 MSSRNRTIYVGNLPGDIRKEVEDLFYKGPVIVDIDLKIPPPPGYAFVEFDPDADD 60
Db 1 MSRENSRTIYVGNLPGDIRKEVEDLFYKGPVIVDIDLKIPPPPGYAF--FEDPDADD 58
Qy 61 AIYGRDGYDFGCLRLVEIAHGGRRFSPVDRYSSYSASR--APRRSDYRVLTGLPP 118
Db 59 AICRGDGYDFGCLRLVEIAHGGRRFSPVDRYSSYSASR--APRRSDYRVLTGLPP 118
Qy 119 SASWQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSDYDMKYAIRKLDATFRNAPSSA 178
Db 118 SASWQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSDYDMKYAIRKLDATFRNAPSSA 178
Qy 179 YIRV-----REYRSRSVSRPDDSKSVRSRSRGRGRCSCSVSSKRSRVSFA 223
Db 178 YIRIICLDAPDCDLSTCSFYDRSVSRSP-----SPV 211
Qy 224 RSISPRPLSRSLYSSVSRSGSLLRAGDWSQSRKSRSRSSNSPV 273
Db 212 -----DERSISRSRTVPSSPSRGRSV-----SKSPRSLSRSPSPV 247

RESULT 13
Q9SPF13 PRELIMINARY; PRT; 261 AA.
AC Q9SPF13;
DT 01-MAY-2000 (TREMELrel. 13; Created)
DT 01-MAY-2000 (TREMELrel. 13; Last sequence update)
DE Splicing factor SR1D.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lazar G., Goddard H.M.;
RA "The Arabidopsis splicing factor SR1 is regulated by alternative
RT splicing.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF173640; AAD52612.1; -
DR HSSP; P19339; 1GXL.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rim; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 261 AA; 28848 MW; A262A497D9BC48BE CRC64;

Query Match 51.3%; Score 745; DB 10; Length 261;
Best Local Similarity 55.4%; Pred. No. 2.9e-58;
Matches 150; Conservative 33; Mismatches 48; Indels 40; Gaps 4;

Qy 1 MSSRNRTIYVGNLPGDIRKEVEDLFYKGPVIVDIDLKIPPPPGYAFVEFDPDADD 60
Db 1 MSSRSRTIYVGNLPGDIRKEVEDLFYKGPVIVDIDLKIPPPPGYAFVEFDPDADD 60
Qy 61 AIYGRDGYDFGCLRLVEIAHGGRRFSPVDRYSSYSASR-----ASRAP 103
Db 61 AIHCRDGYDFGCLRLVEIAHGGRR---SSDTRGSFNGGGRGGRGGRGGRGGRG 117

Qy 104 SRSYRVLTGLPPSASWQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSDYDMKYAIR 163
Db 118 SRSSEFRVLTGLPPSASWQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSDYDMKYAIR 177
Qy 164 RKLDATFRNAPSSAYIRVREYRSRSVSRSPDDSKSVRSRSRGRGRCSCSVSSKRSRVSFA 223
Db 178 KLDATFRNAPSSAYIRVREYRSRSVSRSPDDSKSVRSRSRGRGRCSCSVSSKRSRVSFA 226
Qy 224 RSISPRPLSRSLYSSVSRSGSLLRAGD 254
Db 227 QNL-----HQDLLAPARSQGHRLQEGN 248

RESULT 14
Q8GXSO PRELIMINARY; PRT; 178 AA.
AC Q8GXSO;
DT 01-MAR-2003 (TREMELrel. 23; Created)
DT 01-MAR-2003 (TREMELrel. 23; Last sequence update)
DE Hypothetical protein.
DE Hypothetical protein.
GN AT4G02430/T14P8.21.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA Nakajima M., Enju A., Kamiya A., Natusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RT "Arabidopsis thaliana full-length cDNA.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK118074; BAC42705.1; -
KW Hypothetical protein.
SQ SEQUENCE 178 AA; 19753 MW; DA1E728F07E2B6F3 CRC64;

Query Match 44.8%; Score 649.5; DB 10; Length 178;
Best Local Similarity 69.8%; Pred. No. 5.4e-50;
Matches 125; Conservative 16; Mismatches 21; Indels 17; Gaps 2;

Qy 1 MSSRNRTIYVGNLPGDIRKEVEDLFYKGPVIVDIDLKIPPPPGYAFVEFDPDADD 60
Db 1 MSSRSRTIYVGNLPGDIRKEVEDLFYKGPVIVDIDLKIPPPPGYAFVEFDPDADD 60
Qy 61 AIYGRDGYDFGCLRLVEIAHGGRRFSPVDRYSSYSASR-----SRAPSR 106
Db 61 AIYGRDGYDFGCLRLVEIAHGGRRSSHDA---RGSYSGRGRGGRGGRGGRGGRG 117
Qy 107 SRYRVLTGLPPSASWQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSDYDMKYAIRK 165
Db 118 SRYRVVSLPSSASWQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSDYDMKYAIRK 176

RESULT 15
Q8I290 PRELIMINARY; PRT; 294 AA.
AC Q8I290;
DT 01-NOV-1998 (TREMELrel. 08; Created)
DT 01-NOV-1998 (TREMELrel. 08; Last sequence update)
DE T14P8.21 protein (At4G02430 protein).
GN T14P8.21 OR At4G02430.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
```

Search completed: January 29, 2004, 03:03:34
Job time : 57 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 02:49:20 ; Search time 21 Seconds
(without alignments)
562.130 Million cell updates/sec

Title: US-10-014-927-19

Perfect score: 1451

Sequence: 1 MSSRNRIYVGNLPGDIRK.....RSKRSRSRSPVSPVISG 279

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pap:*

2: /cgn2_6/ptodata/1/1aa/5B.COMB.pap:*

3: /cgn2_6/ptodata/1/1aa/6A.COMB.pap:*

4: /cgn2_6/ptodata/1/1aa/6B.COMB.pap:*

5: /cgn2_6/ptodata/1/1aa/PTUS.COMB.pap:*

6: /cgn2_6/ptodata/1/1aa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	172	11.9	75	1	US-07-881-075-12
2	172	11.9	75	1	US-08-120-827-12
3	172	11.9	75	1	US-08-478-675-12
4	158	10.9	437	2	US-08-935-450-8
5	154.5	10.6	614	5	PCT-US95-03236-21
6	132.5	9.1	579	4	US-09-643-597-348
7	132.5	9.1	579	4	US-09-542-615A-348
8	132.5	9.1	579	4	US-09-608-421B-348
9	131.5	9.1	579	4	US-09-643-597-176
10	131.5	9.1	579	4	US-09-480-884A-176
11	131.5	9.1	579	4	US-09-542-615A-176
12	131.5	9.1	579	4	US-09-608-421B-176
13	130.5	9.0	255	4	US-09-370-838-183
14	129.5	8.9	444	1	US-07-881-075-3
15	129.5	8.9	444	1	US-08-120-827-3
16	129.5	8.9	444	1	US-08-478-675-3
17	127.5	8.8	177	4	US-09-370-838-202
18	126.5	8.7	177	4	US-09-370-838-208
19	126	8.7	380	1	US-07-881-075-51
20	126	8.7	380	1	US-08-120-827-51
21	126	8.7	380	1	US-08-478-675-51
22	125.5	8.6	1850	4	US-09-620-093A-5
23	123.5	8.5	577	3	US-09-261-855-2
24	121	8.3	359	1	US-07-881-075-2
25	121	8.3	359	1	US-08-120-827-2
26	121	8.3	359	1	US-08-478-675-2
27	120	8.3	703	3	US-08-910-925-4

28	115	7.9	299	4	US-09-996-243-213	Sequence 213, Appl
29	114.5	7.9	85	1	US-07-667-276A-9	Sequence 9, Appli
30	114	7.9	68	2	US-08-836-791-5	Sequence 5, Appli
31	113.5	7.8	798	4	US-09-203-453-5	Sequence 5, Appli
32	113.5	7.8	798	4	US-09-900-236-5	Sequence 5, Appli
33	113	7.8	759	1	US-08-676-967-1	Sequence 1, Appli
34	113	7.8	759	1	US-08-676-974-1	Sequence 1, Appli
35	113	7.8	759	2	US-09-098-487-1	Sequence 1, Appli
36	112.5	7.8	655	3	US-09-347-833-4	Sequence 4, Appli
37	112	7.7	217	1	US-08-390-858B-9	Sequence 9, Appli
38	110	7.6	714	2	US-08-990-114-3	Sequence 3, Appli
39	110	7.6	714	4	US-09-241-333-3	Sequence 3, Appli
40	109.5	7.5	743	3	US-08-910-925-3	Sequence 3, Appli
41	108.5	7.5	154	3	US-08-866-928B-5	Sequence 5, Appli
42	108.5	7.5	154	4	US-09-685-836-5	Sequence 5, Appli
43	108.5	7.5	287	1	US-08-146-421-5	Sequence 5, Appli
44	108.5	7.5	471	3	US-08-866-928B-1	Sequence 1, Appli
45	108.5	7.5	471	4	US-09-685-836-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-07-881-075-12
; Sequence 12, Application US/07881075
; Patent No. 5444149
; GENERAL INFORMATION:
; APPLICANT: KEENE, JACK D.
; APPLICANT: KING, PETER H.
; APPLICANT: LEVINE, TODD
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL IN THE
; TITLE OF INVENTION: RECOGNITION, BINDING AND EXPRESSION OF RIBONUCLEIC ACIDS
; TITLE OF INVENTION: INVOLVED IN CELL GROWTH, NEOPLASIA AND IMMUNOREGULATION
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CELON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/881,075
; FILING DATE: 19920511
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5444149man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 714-154-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 75 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-07-881-075-12

Query Match 11.9%; Score 172; DB 1; Length 75;
Best Local Similarity 48.8%; Pred. No. 5.4e-10;
Matches 36; Conservative 11; Mismatches 25; Indels 2; Gaps 1;
QY 9 IYVGNLPGDIRKCEVEDLFVKYGPVIDLKIIPRPPGYAFVEFDPDADDAIYGRDGY 68

Db 2 VYVGNLSSASKHIEGAFKYGPLRV--WVARNPPGFVFEFDRDAEDATRALDGT 59
QY 69 DFDGCLRLRVEIAHG 82
Db 60 RCCGTRIRVEMSSG 73

RESULT 2

US-08-120-827-12
; Sequence 12, Application US/08120827
; Patent No. 5525495
; GENERAL INFORMATION:
; APPLICANT: KEENE, JACK D.
; APPLICANT: KING, PETER H.
; APPLICANT: LEVINE, TODD
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL IN THE
; TITLE OF INVENTION: RECOGNITION, BINDING AND EXPRESSION OF RIBONUCLEIC ACIDS
; TITLE OF INVENTION: INVOLVED IN CELL GROWTH, NEOPLASIA AND IMMUNOREGULATION
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/120,827
; FILING DATE: 15-SEP-1993
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5525495man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 714-158-0 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR

; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 75 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-120-827-12

Query Match 11.9%; Score 172; DB 1; Length 75;
Best Local Similarity 48.6%; Pred. No. 5.4e-10;
Matches 36; Conservative 11; Mismatches 25; Indels 2; Gaps 1;

QY 9 IYVGNLPGDKRCEVEDLFYKYGPVIDLKIIPRPPGYAFVEFDRDADDAIYGRDGY 68
Db 2 VYVGNLSSASKHIEGAFKYGPLRV--WVARNPPGFVFEFDRDAEDATRALDGT 59
QY 69 DFDGCLRLRVEIAHG 82
Db 60 RCCGTRIRVEMSSG 73

RESULT 3

US-08-478-675-12
; Sequence 12, Application US/08478675
; Patent No. 5773246
; GENERAL INFORMATION:
; APPLICANT: KEENE, JACK D.
; APPLICANT: KING, PETER H.

; APPLICANT: LEVINE, TODD
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL IN THE
; TITLE OF INVENTION: RECOGNITION, BINDING AND EXPRESSION OF RIBONUCLEIC ACIDS
; TITLE OF INVENTION: INVOLVED IN CELL GROWTH, NEOPLASIA AND IMMUNOREGULATION
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,675
; FILING DATE: 07-JUN-1996
; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/120,827
; FILING DATE: 15-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5773246man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 714-158-0 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR

; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 75 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-478-675-12

Query Match 11.9%; Score 172; DB 1; Length 75;
Best Local Similarity 48.6%; Pred. No. 5.4e-10;
Matches 36; Conservative 11; Mismatches 25; Indels 2; Gaps 1;

QY 9 IYVGNLPGDKRCEVEDLFYKYGPVIDLKIIPRPPGYAFVEFDRDADDAIYGRDGY 68
Db 2 VYVGNLSSASKHIEGAFKYGPLRV--WVARNPPGFVFEFDRDAEDATRALDGT 59
QY 69 DFDGCLRLRVEIAHG 82
Db 60 RCCGTRIRVEMSSG 73

RESULT 4

US-08-935-450-8
; Sequence 8, Application US/08935450
; Patent No. 5977311
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Yang, Meijia
; APPLICANT: Schulz, Vincent
; TITLE OF INVENTION: 53BP2 COMPLEXES
; FILE REFERENCE: 7934-054
; CURRENT APPLICATION NUMBER: US/08/935,450
; CURRENT FILING DATE: 1997-09-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 8
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-935-450-8


```

Query Match      10.9%; Score 158; DB 2; Length 437;
Best Local Similarity 26.3%; Pred. No. 1.8e-07;
Matches 88; Conservative 29; Mismatches 111; Indels 106; Gaps 15;

QY 9 IYVGNLPGDIRKCEVEDLFYKYGPIVDIDL---KIPRPPGYAFVEFEDPRDADDAIYGR 65
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
10 LFIGGLTETNEKALEAVEFGYGRIVEVLLMKDRETNKSRGFAFVTPESPADAKDAARDM 69
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
66 DQYDFDGCRLRVEIA-----HGRFPSPVDR----- 92
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
70 NGKSLDGKAIKVEQATKPSFESGRGPPPPRSGRGLRGRGGSGGTRGPPSRGGHM 129
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
93 ----YSSYSAS-----RAPSRSDYRLVLTGLPP--SASQDLKDHWRKAGDVCF 137
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
130 DDGGSXMFNMFSSRGPLVTKRGPSPRG-----GPPPKSAPSGPVRRSSSGMGGRAPV 183
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
138 S-----EYFPDRK-----GMSGVVDYGNVD-----DMKYAI 163
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
184 SRGRDSYGGPREPLPSRROVLSPRDDGYSTKDSYSSRDYPSRDRDYAPPRDYTY 243
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
164 RKLDATERFNAFSS-----AYIRVREYSRSVSRSPDDG-KSY---RSRSGRGPSCS 212
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
244 RDYGHSSRDDYPSREYSDRDYGRDRDYSRHPGSGSYRDSVESYGNRSRSPPTRGPPPS 303
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
213 YSSKSR--SVSPARSISPSR-PLSRSS-LYSS 242
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
304 YGSSRYDDYSSRDYGGSGRDSYSSSRSDLYSS 337

```

```

RESULT 5
PCT-US95-03236-21
; Sequence 21, Application PC/TUS9503236
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: Methods to Diagnose and Treat HIV-1
; TITLE OF INVENTION: Infection
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03236
; FILING DATE: 13-MAR-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Imbra, Richard J.
; REGISTRATION NUMBER: 37,643
; REFERENCE/DOCKET NUMBER: FP-SI 1394
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 614 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-03236-21

```

```

Query Match      10.6%; Score 154.5; DB 5; Length 614;
Best Local Similarity 26.6%; Pred. No. 6.6e-07;
Matches 66; Conservative 19; Mismatches 100; Indels 63; Gaps 6;

```

```

QY 7 RTIYVGNLPGDIRKCEVEDLFYKYGPIVDIDL---KIPRPPGYAFVEFEDPRDADDAIY 63
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
280 KTLFVARVNVDTTESKLRREFEVYGPILRHVYVSRSGKPGVAFIEYEHEDMHSAYK 339
QY 64 GRGDFDGCRLRVEIAHGRFPSPVSDRYSSYSASRAPSRSDYRLVLTGLPPSASWQ 123
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
340 HADGCKIDGRRLVDVERGR-----TVKGRPRRLGG 371
QY 124 DLKDHMRKAGDVCFSEVPDRKMGSGVVDYSNYDD-----MKYAIRKLDATERFNAFSS 177
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
372 GLGGRGAGADVNRH-----SGRDDTSYRDERPGSPDUPHRDRDSD----- 413
QY 178 AYIRVREYSRSVSRSPDDSKSYRSRSGRGPSCSYSSKSRSVSPARSISPSRPLSRSR 237
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
414 -----RERERERSREDKE---RERRSRSDRRRRSRSDKEERRSERESKDKDRR 465
QY 238 SLYSSVSR 245
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
466 KRSSRSR 473

RESULT 6
US-09-643-597-348
; Sequence 348, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-643-597-348

Query Match      9.1%; Score 132.5; DB 4; Length 579;
Best Local Similarity 23.5%; Pred. No. 9.5e-05;
Matches 64; Conservative 45; Mismatches 116; Indels 47; Gaps 12;

QY 9 IYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPRPP-----GYAFVEFEDPRDADDAIY 63
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
4 LVIGNLSENAAPSDLESIFK-----DAKIPVSGPFLVKGTGYAFVDCPDESWALKAE 55
QY 64 GRGDFDGCRLRVEIAHGRFPSPVSDRYSSYSASRAPSRSDYRLVLTGLPPSASWQ 123
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
56 ALSG-----KIEL-HG-----KPIEVHS-----VPKRQIRKQLQIRNIPPHLOWE 95
QY 124 DLKDHMRKAGDV--CFSEVPDRKMGSGVVDYSNYDDMKYAIRKLDATERFNAFSSAYI 180
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
96 VLDSLVLQYGVVESC-EQVNTDSETAVVNTVSSKQDQARQALDKLNGFOLENTLXVAI 154
QY 181 RVREYSRSVSRSPDDSKSYRSRSGRGPSCSYSSKSRSVS-PARSISPSRPLSRSRSL 239
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
155 PDETAQQNPLOQPRGRGLGQGRSGRQSGPSGVSKQKPCDPLRLLVF-----TQFV 207
QY 240 YSSVSRSGSLLAGDWISQSRKSRSRSRNS 271
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
208 GAIIGKEGATIR--NITKQTSKIDVHRKENA 237

```

```
/ LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-606-421B-348
```

```
Query Match      9.1%; Score 132.5; DB 4; Length 579;  
Best Local Similarity 23.5%; Pred. No. 9.5e-05;  
Matches 64; Conservative 45; Mismatches 116; Indels 47; Gaps 12;
```

```
QY   9 IYVGNLPGDIRKCEVEDLFKYGPVIDIDLKIIPRPP-----GYAFVEFEDPRDADDAYI 63  
    :|::||:|:|:|:|:|:|:|:|:|:|:|:~::~:~::~:~::~:  
DB   4 LYIGNLSENAAPSDESIFK-----DAKIPVGPFVLVTGTFYAVDCPDSEWALKAE 55  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
```

```
QY   64 GRGDYDFDCRLRVETIAHGRFRPSVDRYSSYSASRAPRRSDYRVLTGLPSPASWQ 123  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:  
DB   56 ALSG-----KIEL-HG----KPIVEHS-----VPKRQRIRKLQIRNIPEHLQWE 95  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
```

```
QY   124 DLKDHRMKAADV--CFSEVFDPDRKGMSGVDDYNVKVAIRKLDATERN-AFSXAI 180  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:  
DB   96 VLDSLLVOYGVVESC-EQNWTSETAUVNNTVSQQARQLDKNGLENFLTKVAYI 154  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:  
QY   181 RVREYESRSVRSPDDSKSYRSRSGRPSCSVSCKSRSVS-PARSISPRSPLSRSL 239  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:  
DB   155 PDETAACNQPLQPGRRGRLGQGSSROGPSVSKOKPCDLPLRLLP-----TQFV 207  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
```

```
QY   240 YGSVRSGLLRAGDWISOCSKSRSRSNS 271  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:  
DB   208 GAIGKEGATR-NITKTQS KIDVHRKENA 237  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
```

```
RESULT 9  
US-09-643-597-176  
; Sequence 176, Application US/09643597  
; Patent No. 6426072  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455CII  
; CURRENT APPLICATION NUMBER: US/09/643,597  
; CURRENT FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 369  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 176  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-643-597-176
```

```
Query Match      9.1%; Score 131.5; DB 4; Length 579;  
Best Local Similarity 23.5%; Pred. No. 0.00012;  
Matches 64; Conservative 45; Mismatches 116; Indels 47; Gaps 12;
```

```
QY   9 IYVGNLPGDIRKCEVEDLFKYGPVIDIDLKIIPRPP-----GYAFVEFEDPRDADDAYI 63  
    :|::||:|:|:|:|:|:|:|:~::~:~::~:~::~:  
DB   4 LYIGNLSENAAPSDESIFK-----DAKIPVGPFVLVTGTFYAVDCPDSEWALKAE 55  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
```

```
QY   64 GRGDYDFDCRLRVETIAHGRFRPSVDRYSSYSASRAPRRSDYRVLTGLPSPASWQ 123  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:  
DB   56 ALSG-----KIEL-HG----KPIVEHS-----VPKRQRIRKLQIRNIPEHLQWE 95  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
```

```
QY   124 DLKDHRMKAADV--CFSEVFDPDRKGMSGVDDYNVKVAIRKLDATERN-AFSXAI 180  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:  
DB   96 VLDSLLVOYGVVESC-EQNWTSETAUVNNTVSQQARQLDKNGLENFLTKVAYI 154  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:  
QY   181 RVREYESRSVRSPDDSKSYRSRSGRPSCSVSCKSRSVS-PARSISPRSPLSRSL 239  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:  
DB   155 PDETAACNQPLQPGRRGRLGQGSSROGPSVSKOKPCDLPLRLLP-----TQFV 207  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
```

```
QY   240 YGSVRSGLLRAGDWISOCSKSRSRSNS 271  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:  
DB   208 GAIGKEGATR-NITKTQS KIDVHRKENA 237  
    :|::||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
```

```
US-09-606-421B-348  
; Sequence 348, Application US/09606421B  
; Patent No. 6531315  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosket, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C9  
; CURRENT APPLICATION NUMBER: US/09/606,421B  
; CURRENT FILING DATE: 2000-06-28  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 348
```

Db 96 VLDSLIVQYGVESC-EQVNTDSETAVVNTYSSKDQARQALDKLNGFLENFTLKVAYI 154
QY 181 RVREYERSVSRPDDSKSYRSRSGRGPCSYSSKSRSVS-PARSISPRSRPLSRSL 239
Db 155 PDMAAQQNPLOQPRGRGLGQSGRSRQSGSVSKQKPCDPLRLVLP-----TQFV 207
QY 240 YSVSRSGSLLRAGDWISQSRKSRSRSRNS 271
Db 208 GAIIGEGATIR--NITKQTSKIDVHRKENA 237

RESULT 10
US-09-480-884A-176
; Sequence 176, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-480-884A-176

Query Match 9.1%; Score 131.5; DB 4; Length 579;
Best Local Similarity 23.5%; Pred. No. 0.00012;
Matches 64; Conservative 45; Mismatches 116; Indels 47; Gaps 12;
QY 9 IYVGNLPGIRKCEVEDLFYKGPVIDIDLKIPRPP-----GYAFVEFEDPRDADDAIY 63
Db 4 LYIGNLSENAAPSDLESIFK-----DAKIPVSGPFLVKTGYAFVDCPDESALKAE 55
QY 64 GRDGYDFDGCRLRVEIAHGGRFRFSPVDYSSYSASRAPSRSDYRVLTGLPPSASWQ 123
Db 56 ALSG-----KIEL-HG-----KPIEVHS-----VFKQRIKQLQIRNIPPHLOWE 95
QY 124 DLKDHMRKAGDV--CFSEVFPDRKMGSGVVDYSDMKYAIRKLDATFERN-AFSSAYI 180
Db 96 VLDSLIVQYGVESC-EQVNTDSETAVVNTYSSKDQARQALDKLNGFLENFTLKVAYI 154
QY 181 RVREYERSVSRPDDSKSYRSRSGRGPCSYSSKSRSVS-PARSISPRSRPLSRSL 239
Db 56 ALSG-----KIEL-HG-----KPIEVHS-----VFKQRIKQLQIRNIPPHLOWE 95
QY 124 DLKDHMRKAGDV--CFSEVFPDRKMGSGVVDYSDMKYAIRKLDATFERN-AFSSAYI 180
Db 96 VLDSLIVQYGVESC-EQVNTDSETAVVNTYSSKDQARQALDKLNGFLENFTLKVAYI 154
QY 181 RVREYERSVSRPDDSKSYRSRSGRGPCSYSSKSRSVS-PARSISPRSRPLSRSL 239
Db 155 PDMAAQQNPLOQPRGRGLGQSGRSRQSGSVSKQKPCDPLRLVLP-----TQFV 207
QY 240 YSVSRSGSLLRAGDWISQSRKSRSRSRNS 271
Db 208 GAIIGEGATIR--NITKQTSKIDVHRKENA 237

RESULT 11
US-09-542-615A-176
; Sequence 176, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C8

; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-542-615A-176
Query Match 9.1%; Score 131.5; DB 4; Length 579;
Best Local Similarity 23.5%; Pred. No. 0.00012;
Matches 64; Conservative 45; Mismatches 116; Indels 47; Gaps 12;
QY 9 IYVGNLPGIRKCEVEDLFYKGPVIDIDLKIPRPP-----GYAFVEFEDPRDADDAIY 63
Db 4 LYIGNLSENAAPSDLESIFK-----DAKIPVSGPFLVKTGYAFVDCPDESALKAE 55
QY 64 GRDGYDFDGCRLRVEIAHGGRFRFSPVDYSSYSASRAPSRSDYRVLTGLPPSASWQ 123
Db 56 ALSG-----KIEL-HG-----KPIEVHS-----VFKQRIKQLQIRNIPPHLOWE 95
QY 124 DLKDHMRKAGDV--CFSEVFPDRKMGSGVVDYSDMKYAIRKLDATFERN-AFSSAYI 180
Db 96 VLDSLIVQYGVESC-EQVNTDSETAVVNTYSSKDQARQALDKLNGFLENFTLKVAYI 154
QY 181 RVREYERSVSRPDDSKSYRSRSGRGPCSYSSKSRSVS-PARSISPRSRPLSRSL 239
Db 155 PDMAAQQNPLOQPRGRGLGQSGRSRQSGSVSKQKPCDPLRLVLP-----TQFV 207
QY 240 YSVSRSGSLLRAGDWISQSRKSRSRSRNS 271
Db 208 GAIIGEGATIR--NITKQTSKIDVHRKENA 237

RESULT 12
US-09-606-421B-176
; Sequence 176, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-176

Query Match 9.1%; Score 131.5; DB 4; Length 579;
Best Local Similarity 23.5%; Pred. No. 0.00012;
Matches 64; Conservative 45; Mismatches 116; Indels 47; Gaps 12;
QY 9 IYVGNLPGIRKCEVEDLFYKGPVIDIDLKIPRPP-----GYAFVEFEDPRDADDAIY 63
Db 4 LYIGNLSENAAPSDLESIFK-----DAKIPVSGPFLVKTGYAFVDCPDESALKAE 55
QY 64 GRDGYDFDGCRLRVEIAHGGRFRFSPVDYSSYSASRAPSRSDYRVLTGLPPSASWQ 123
Db 56 ALSG-----KIEL-HG-----KPIEVHS-----VFKQRIKQLQIRNIPPHLOWE 95

RESULT 15
 US-08-120-827-3
 / Sequence 3, Application US/08120827
 / Patent No. 5525495
 /
 / GENERAL INFORMATION:
 /
 / APPLICANT: KEENE, JACK D.
 /
 / APPLICANT: KING, PETER H.
 /
 / APPLICANT: LEVINE, TODD
 /
 / TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL IN THE
 / TITLE OF INVENTION: RECOGNITION, BINDING AND EXPRESSION OF
 / TITLE OF INVENTION: INVOLVED IN CELL GROWTH, NEOPLASIA AND
 / NUMBER OF SEQUENCES: 101
 /
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
 / ADDRESSEE: P.C.
 / STREET: 1755 Jefferson Davis Highway, Fourth Floor
 / City: Arlington
 / STATE: Virginia
 / COUNTRY: U.S.A.
 / ZIP: 22202
 /
 / COMPUTER READABLE FORM:
 /

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/120,827
FILING DATE: 15-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Obolon, No. 5525495man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 714-158-0 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)413-3000
TELEFAX: (703)413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 444 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: Protein
US-08-120-827-3

Query Match 8.9%; Score 129.5; DB 1; Length 444;
Best Local Similarity 21.5%; Pred. No. 0.00013;
Matches 58; Conservative 47; Mismatches 122; Indels 43; Gaps 8;
QY 9 IYVGNLPGDIRKCEVEDLFYKYGPVIDL---KIPPEPPGYAFVEPEDPRDADDAIYGR 65
Db 112 LIVNLPDTMSQDEIRSLFVSGVESCKLRDKVTGOSLGGYGVYKQEDAEKAINAL 171
QY 66 DGYDFDGCRLRVEIAHGRRFPSPVDYSSYSASRAPRRSDYRVLVLTGLPPSASWQDL 125
Db 172 NGLRLQNTIKVSIAR-----PSESIGKA-----NLYVSLPKNTQSDLL 212
QY 126 KDHMRKAGDVCFSEVFPDR-----KMGSGVVDYSNYD---DMKYAIRKLDATFERNAPS 176
Db 213 ESLFSPYCKIITSRLCDNITDEHAAGLSKGVGFIRDPQFEADRAIKELNGTTPKNSTE 272
QY 177 SAYIRVREYSRSRSPDDSKYSRSRSGRP---SCSYSSKRSRVSPPARSISPRSRPL 233
Db 273 PITVKFANNPS-SNKNMSQPLAAAYIAPQNTGRGFAFPANAAAGAAAAAAAIHPNAGRY 331
QY 234 SRSRLYSVSR-----SGSLLRAGDW 255
Db 332 SSVISRSPLTSDLTNGMIQNTIASGW 361

Search completed: January 29, 2004, 02:57:34
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 02:50:21 ; Search time 243 Seconds
(without alignments)
237.364 Million cell updates

Title: US-10-014-927-19

Perfect score: 1451
Sequence: 1 MSSRWNRITTVGNLPGDIRK.....RSKSRSRSRNSPVPVTSG 279

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 777136 seqs, 206736638 residues

Total number of hits satisfying chosen parameters: 777136

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:**

```

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/FCI_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/FCIUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US108_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1451	100.0	279	14	US-10-014-927-19	Sequence 19, Appl
2	582.5	40.1	248	14	US-10-014-927-22	Sequence 22, Appl
3	517	35.6	253	10	US-09-925-300-1601	Sequence 1601, Ap
4	517	35.6	254	12	US-10-264-049-2448	Sequence 2448, Ap
5	277	19.1	306	9	US-09-925-297-601	Sequence 601, App
6	275.5	19.0	94	12	US-09-864-408A-344	Sequence 344, App
7	195.5	13.5	155	9	US-09-925-302-759	Sequence 759, App
8	177.5	12.2	137	12	US-10-104-047-2470	Sequence 2470, Ap
9	158	10.9	364	15	US-10-153-668-472	Sequence 472, App
10	156.5	10.8	112	12	US-10-264-049-2584	Sequence 2584, Ap
11	153	10.5	223	12	US-10-264-049-2995	Sequence 2995, Ap
12	152	10.4	366	15	US-10-153-668-228	Sequence 228, App
13	151.5	10.4	366	15	US-10-177-293-423	Sequence 423, App
14	146.5	10.1	250	12	US-10-197-666A-78	Sequence 78, Appl
15	146.5	10.1	252	12	US-10-108-260A-4815	Sequence 4815, Ap

ALIGNMENTS

RESULT 1

```

US-10-014-927-19
; Sequence 19, Application US/10014927
; Publication No. US2002011580A1
; GENERAL INFORMATION:
; APPLICANT: Barta, Andrea
; APPLICANT: Lopato, Sergiy
; APPLICANT: Kalyra, Maria
; APPLICANT: Dörner, Silke
; TITLE OF INVENTION: Splice Factor
; FILE REFERENCE: SONN-013US
; CURRENT APPLICATION NUMBER: US/10/014, 927
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: PCT/AT00/00100
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: A 727/99
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of the un
; -OTHER INFORMATION: atSRP30
US-10-014-927-19

```

Query Match 100.0%: Score 1451: DB 14: Length 279:

Query Match	100.0%;	score 1431; DB 1431
Best Local Similarity	100.0%;	Pred. No. 1.4e-128;

Best local similarity 100.0%, freq. no. 1.48-120,
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	

1 MSSRWNRTIYVGNLPGDIRKCEVEDLFYKYGPVDIDLKIPRPPGYAFVEFEDPRDADD 60

1 MSSRWNRITYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPRPPGYAFVEFEDPRDADD 60

[illegible]

61 AIYGRDGYDFDGCRLRVEIAHGGRFFSPVDRYSSSYASRAPSRKSDYRLVLTGLPFA 12

Db 61 AIYGRDGYDFDGCRLRVEIAHGGRRFSPVDRYSSYSASRAPRRSDYRVLVTGLPPSA 120
QY 121 SQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSDYDMKYAIRKLDATFERNAPSSAYI 180
Db 121 SQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSDYDMKYAIRKLDATFERNAPSSAYI 180
QY 181 RVREYERSVSRSPDDSKSVRSRSGRSPGSCYSKRSVSPARSISPRSRPLSRGRSLY 240
Db 181 RVREYERSVSRSPDDSKSVRSRSGRSPGSCYSKRSVSPARSISPRSRPLSRGRSLY 240
QY 241 SSVSRSGSLLFAGDWISQSRKSRSGRSGNSPVSFVIGS 279
Db 241 SSVSRSGSLLFAGDWISQSRKSRSGRSGNSPVSFVIGS 279

RESULT 2
US-10-014-927-22
; Sequence 22, Application US/10014927
; Publication No. US20020115180A1
; GENERAL INFORMATION:
; APPLICANT: Barta, Andrea
; APPLICANT: Lopato, Sergio
; APPLICANT: Kalyna, Maria
; APPLICANT: Dörner, Silke
; TITLE OF INVENTION: Splice Factor
; FILE REFERENCE: SONN:01305
; CURRENT APPLICATION NUMBER: US/10/014, 927
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: PCT/AT00/00100
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: A. 727/99
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-014-927-22

Query Match 40.1%; Score 582.5; DB 14; Length 248;
Best Local Similarity 53.6%; Pred. No. 9.5e-47;
Matches 134; Conservative 28; Mismatches 47; Indels 41; Gaps 8;
QY 9 IYVGNLPGDIRKCEVEDLFYKYGPVVDIDLKIPRPPGYAFVFEFDPDADDAIYGRDGY 68
Db 18 IYVGNLPGDIRKCEVEDLFYKYGPVVDIDLKIPRPPGYAFVFEFDPDADDAIYGRDGY 77
QY 69 DFDGCRRLRVEIAHGGRRFSPVDRYSSYSASRAPRRSDYRVLVTGLPPSA 111
Db 78 DYDGYRLRVEFPGRSGRGTGRGGGGGGGAPGRGYGPD-----PSRRSENRY 123
QY 112 LVTGLPPSASQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSDYDMKYAIRKLDATF 171
Db 124 VVSGLPSPGSGQDLKDHMRKAGDVCYADVVDG---TGVEFVRKEDMTYAVAKLDNTKF 180
QY 172 R-NAPSSAYIRVREYERSVSRSPDDSKSVRSRSGRSPGSCYSKRSVSPARSISPR 229
Db 181 RSHGEETAYIRKVDGPRSPSYGRSGRS-RSRSGRSRS---NSRSGRSYPRSRGSPR 236
QY 230 SRPL-SRSGRS 238
Db 237 YSPRHSRSGRS 246

RESULT 3
US-09-925-300-1601
; Sequence 1601, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1601
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1601

Query Match 35.6%; Score 517; DB 10; Length 253;
Best Local Similarity 52.5%; Pred. No. 1.5e-40;
Matches 115; Conservative 30; Mismatches 38; Indels 36; Gaps 8;
QY 9 IYVGNLPGDIRKCEVEDLFYKYGPVVDIDLK-----IPRPPGYAFVFEFDPDADDAIY 63
Db 48 IYVGNLPGDIRKCEVEDLFYKYGPVVDIDLK-----IPRPPGYAFVFEFDPDADDAIY 102
QY 64 GEDGYDFDGCRLRVEI--AHGRRFSPVDRYSSYSASRAPRRSDYRVLVTGLPPSA 121
Db 103 GRNGDYGCRLRVEFPRTYGGRGWPRGGR-----NGPTRSDPRVLVSLPPSGS 155
QY 122 WQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSDYDMKYAIRKLDATFERNAPSSAYI 180
Db 156 WQDLKDHMRKAGDVCYADVVDGVM---VEYLRKEDMEYALRKLDDTKFRSHEGETSYI 212
QY 181 RVREYERSVSRSPDDSKSY---RSRSGRSGPSCYSK 216
Db 213 RV-----YPERSTSYGYSRSGRSGRSDPYQSR 241

RESULT 4
US-10-264-049-2448
; Sequence 2448, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2448
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-2448

Query Match 35.6%; Score 517; DB 12; Length 254;
Best Local Similarity 52.5%; Pred. No. 1.5e-40;
Matches 115; Conservative 30; Mismatches 38; Indels 36; Gaps 8;
QY 9 IYVGNLPGDIRKCEVEDLFYKYGPVVDIDLK-----IPRPPGYAFVFEFDPDADDAIY 63
Db 49 IYVGNLPGDIRKCEVEDLFYKYGPVVDIDLK-----IPRPPGYAFVFEFDPDADDAIY 103
QY 64 GEDGYDFDGCRLRVEI--AHGRRFSPVDRYSSYSASRAPRRSDYRVLVTGLPPSA 121
Db 104 GRNGDYGCRLRVEFPRTYGGRGWPRGGR-----NGPTRSDPRVLVSLPPSGS 156
QY 122 WQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSDYDMKYAIRKLDATFERNAPSSAYI 180
Db 157 WQDLKDHMRKAGDVCYADVVDGVM---VEYLRKEDMEYALRKLDDTKFRSHEGETSYI 213

QY 181 RVREYRSVSRSPDSSKY---RSTRSRSGPSCSYSK 216
 Db 214 RV-----YPERSTSYGYSRSGSGRGRDSPYQSR 242

RESULT 5

US-09-925-297-601
 ; Sequence 601, Application US/09925297
 ; Patent No. US20020081659A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA105
 ; CURRENT APPLICATION NUMBER: US/09/925,297
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05989
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 928
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 601
 ; LENGTH: 306
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-925-297-601

Query Match 19.1%; Score 277; DB 9; Length 306;
 Best Local Similarity 38.8%; Pred. No. 8.2e-18;
 Matches 85; Conservative 27; Mismatches 65; Indels 42; Gaps 7;

QY 93 YSSSYASAPRRSDYRVLTGLPSPASWDLKDHMKAGDVCFSEVFPDRKMGSGVVD 152
 Db 20 YRRSGDKYGPTRTEYLLIVENLSSRCWDLKDYMFQAGEVYADAHKGRKN-EGVIE 78

QY 153 YSNYDMKYAIRKLDATFR-----NAPSSAYIRVEYSRVSRS----- 193
 Db 79 FVSYDMKRALEKLDGTEVNGRKILVDEKPGSRRRRSYSRSH-SRSRSHRSHKSR 137

QY 194 --PDDSKYRSRSGPSCSY-SKRSVSPARSISPR-----SRPLSRSLYSVS 244
 Db 138 SRSGSKSHSKRSRSGRSRSGRSRSGRSRSGRSRSGRSRSGRSRSGRSRSGRS 197

QY 245 RGSLLRAGDWI-----SQRSKRSRGRSRS 271
 Db 198 RSKSKDAEEKIQNDNVGPKRSRSPSRHKSRSRSGRS 236

RESULT 6

US-09-864-408A-344
 ; Sequence 344, Application US/09864408A
 ; Publication No. US20040009474A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Leach, Martin D.
 ; APPLICANT: Shimkets, Richard A.
 ; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encoded by the Same
 ; FILE REFERENCE: 21402-012
 ; CURRENT APPLICATION NUMBER: US/09/864,408A
 ; CURRENT FILING DATE: 2001-05-24
 ; PRIOR APPLICATION NUMBER: 60/206,690
 ; PRIOR FILING DATE: 2000-05-24
 ; NUMBER OF SEQ ID NOS: 9068
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 344
 ; LENGTH: 94
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(1)
 ; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
 US-09-864-408A-344

Query Match 19.0%; Score 275.5; DB 12; Length 94;
 Best Local Similarity 50.0%; Pred. No. 2.3e-18;
 Matches 68; Conservative 8; Mismatches 17; Indels 43; Gaps 4;

QY 134 DVCSEVFPDRKMGSGVVDYDNDKYYAIRKLDATFRNAPSSAYIRVEYSRVSRS 193
 Db 2 DVCSEVYREGGTGIVDTYDNDKYYAIRKLDATFRNAPGRAYIRKEY----- 53

QY 194 PDDSKYRSRSGPSCSYSKRSVSPARSISPRSLRSLRSLRSLRSLRSLRSLRSLR 253
 Db 54 --NGKGRSYRSRSPSRSY-SKRSRPSK-----SPRTR----- 84

QY 254 DWISQSRKSRSRSR 269
 Db 85 -----RSSSRSRSR 94

RESULT 7

US-09-925-302-759
 ; Sequence 759, Application US/09925302
 ; Patent No. US20020044941A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA104
 ; CURRENT APPLICATION NUMBER: US/09/925,302
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05918
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 896
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 759
 ; LENGTH: 155
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: SITE
 ; LOCATION: (147)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 US-09-925-302-759

Query Match 13.5%; Score 195.5; DB 9; Length 155;
 Best Local Similarity 37.1%; Pred. No. 1.6e-10;
 Matches 46; Conservative 18; Mismatches 33; Indels 27; Gaps 3;

QY 9 IYVGNLPGDIRKCEVEDLFYKGYPIVDIDLKIPRPPGVAFVEFEDPRDADDAIYGRDGY 68
 Db 43 VYVGNLGNNGKTELEAFYGYGLRSV--WVARNPPGFAFVEFEDPRDADAVRELDGR 100

QY 69 DFDGCRRLRVEIAHGRRFPSPVDYSSYSASRAPSRRSDYRVLTGLPSPASWDLKDH 128
 Db 101 TLGCRVRLVSLNGEKR-----SRNRGP-----PPSWGRRPRDDY 135

QY 129 MRKA 132
 Db 136 RRS 139

RESULT 8

US-10-104-047-2470
 ; Sequence 2470, Application US/10104047
 ; Publication No. US20030236392A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: No. US20030236392A1 full length cDNA
 ; FILE REFERENCE: H1-A0105
 ; CURRENT APPLICATION NUMBER: US/10/104,047
 ; CURRENT FILING DATE: 2002-03-25
 ; PRIOR APPLICATION NUMBER:
 ; PRIOR FILING DATE:


```

Db      95  KUHGVNINVE-----ASKUKSTST-KUHGVNISTCTNKELRK 133
Qy      129  MKKAGVCFSEVFPDRKMGSGVDDMKYAIRKDATEPRN-----AFSAYIRVR 183
Db      134  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      134  FBEYGPETCDIKV-----YAFVHMRAEDAVEAIRGLDNTFQGRMHVQLSTSLRTA 189
Qy      184  EYESRSVRSRSPDCKSYRSRSR--SRGPGSCSYSKSR 218
Db      190  PWTGR-----PWLILVMVERALVORVPSRSGYSGCR 221

```

RESULT 12

```

US-10-153-568-228
; Sequence 228, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi
; APPLICANT: MATSUDA, Aki
; APPLICANT: MURAMATSU, Shuji
; APPLICANT: ISHIZAWA, Kenya
; TITLE OF INVENTION: STATE Activating Gene
; FILE REFERENCE: 1254-0207P
; CURRENT APPLICATION NUMBER: US/10/153,568
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/293,172
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/316,031
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/328,403
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: JP 2001-157043
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: JP 2001-260681
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: JP 2001-313175
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 488
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 228
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-568-228

```

Query Match	10.5%	Score 152;	DB 15;	Length 366;
Best Local Similarity	24.9%;	Pred. NO. 6.5e-06;		
Matches	48;	Conservative 33;	Mismatches 80;	Indels 32; Gaps 5
Qy	9	IYVGNLGGDIRCEVEDLFYKGPVTD;DLKIPRPQGYAFVEFEDPRDADDAIYGRGY	68	
Db	4	LFIGNLPRATQETRSIFEQGVKLECDI-----IKNYGFVHIEDTAAEDAIRNLHY	58	
Qy	69	DFDQCLRLVEIAHGGRRFSPSVDRYSSYSASRAPSRSDYRVLVTGLPPSPASQDLKH	128	
Db	59	KLHGVNINVE-----ASQKSKTST-KLHVGNISPTCTNKELRAK	97	
Qy	129	MRKAGVCFSEFPDRKGMGVVDYSNDDMKYAIRKLDATEPERNAFSSAYRVREYSR	188	
Db	98	FHEYGVFIECDIKVQ-----YAFVHMERAEADVAEIRGLDNTFEGG--KRMHVQLSTSLRL	151	
Qy	189	SVSRSPDDSKSYR	201	
Db	152	TAPGMGDQSGCYR	164	

RESULT 13

RESUME 13
US-10-177-293-423
; Sequence 423. Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei

; APPLICANT: Gannavarpu, Manjula
 ; APPLICANT: Kamatkar, Shubhangi
 ; APPLICANT: Mertens, Maureen
 ; APPLICANT: Myer, Vic
 ; APPLICANT: Wang, Youzhen
 ; APPLICANT: Xu, Yongyao
 ; APPLICANT: Hoersch, Sebastian
 ; APPLICANT: Monahan, John
 ; APPLICANT: Meyers, Rachel E.
 ; APPLICANT: Bast Jr., Robert C.
 ; APPLICANT: Hortobagyi, Gabriel N.
 ; APPLICANT: Pusztai, Lajos
 ; APPLICANT: Meric, Funda
 ; APPLICANT: Sahin, Aysegul
 ; APPLICANT: Mills, Gordon B.
 ; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
 ; PREVENTION, AND THERAPY OF BREAST CANCER
 ; FILE REFERENCE: MRI-038
 ; CURRENT APPLICATION NUMBER: US/10/177,293
 ; CURRENT FILING DATE: 2002-06-21
 ; PRIOR APPLICATION NUMBER: US 60/299,887
 ; PRIOR FILING DATE: 2001-06-21
 ; PRIOR APPLICATION NUMBER: US 60/301,572
 ; PRIOR FILING DATE: 2001-06-27
 ; PRIOR APPLICATION NUMBER: US 60/306,501
 ; PRIOR FILING DATE: 2001-07-18
 ; PRIOR APPLICATION NUMBER: US 60/325,002
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US 60/362,585
 ; PRIOR FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
 ; PRIOR FILING DATE: 2002-05-14
 ; NUMBER OF SEQ ID NOS: 506
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 423
 ; LENGTH: 3664
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-177-293-423

Query Match	10.4%;	Score 151.5;	DB 15;	Length 3664;
Best Local Similarity	26.6%;	Pred. No. 0.00017;		
Matches	85;	Conservative 38;	Mismatches 128;	Indels 69; Gaps 13;

Qy	4	RWNETVGNLPDGIKCEVEDL.FYKXGPIVDLDKLIPRPPG----	YAFVEFEDPRDAD	59
Db	3	RETRHLWGNLPENVEEKIIEHFYGRVSEV--XILPKRGSEGGVAAVFDFDIKSAQ	60	
Qy	60	DAIYG-----RDCYDFDGCRLRVEIAHGRRRSPSVDRVSSYSASRAPSRSDYR	110	
Db	61	KAHNSVNMKGDRDLRLRTDYNFG--TIPSAARGLDDTVSIASRSREVSFGFGGGGPAY-	116	
Qy	111	VLVTGLPPGASWCD-----LKDHMKAGVCFSEVPFPDRKMGMSVVDYNYDD----	158	
Db	117	----GPPPSLHAREGRYERRLDGASDNRRAYE--HSAVGHHERGTGGFDRTHYDQDY	170	
Qy	159	-----MKYARKLDATEF-----RNAFSSAIRVREYSRSVSRSPDD	196	
Db	171	RDPRERTLQHLGYAGRSRSPNPFDAHDPRYEPAREQFTLPGVVHMDIYRDDITREVRG	230	
Qy	197	SKGYS--RSRSGPSCSYSSKRSVPARSISPRGRPLSRSRSLYSSVSRSGLLRAGD	254	
Db	231	RRPERYQHRSRSP---HSSQBNQSPQLASQASRP-TRSPGSGSRSRSSS----	SD	282
Qy	255	WISQSRKSRSRGRSRNSPV	274	
Db	283	SISSSSTSDSDSDSSSSS	302	

RESULT 14

US-10-197-666A-78
; Sequence 78, Application US/10197666A
; Publication No. US20030092037A1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2004, 20:16:44 ; Search time 64 Seconds

(without alignments)
1531.048 Million cell updates/sec

Title: US-10-014-927-19MOD_COPY_1_222

Perfect score: 222

Sequence: 1 MRRNRRTIVGNLPDIRK.....RRSRGRPCSYSSKRSRVSYP 222

Scoring table: UNITARY2

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/BAUM927/runat.04022004.131334.2726/app.query.fasta_1.391
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=unitary2 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=BAUM927 @CEN 1 56 @runat.04022004.131334.2726 -NCPUS=6 -ICPU=3 -NO_WMAP
-LARGEQUERY -NEG SCORES=0 -WAIT -DISPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Issued Patents NA:*
- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

WDC for detected items + 6 WDS

Result No.	Score	Query Match	Length	DB ID	Description
1	48	21.6	284	4	US-09-313-294A-5971
2	48	21.6	888	4	Sequence 5971, Ap
3	48	21.6	1824	4	Sequence 12914, A
C 4	48	21.6	1824	4	Sequence 12570, A
C 5	48	21.6	4403765	3	US-09-352-991A-12570
C 6	47	21.2	955	4	Sequence 2, Appli
7	47	21.2	1336	4	Sequence 524, Appli
8	47	21.2	2508	4	Sequence 12914, A
C 9	47	21.2	11282	4	Sequence 28, Appl
C 10	46	20.7	771	4	Sequence 793, Appl
11	46	20.7	888	4	Sequence 3, Appli
12	46	20.7	1420	4	Sequence 8080, Ap
					Sequence 8235, Ap
					Sequence 23, Appli
					Sequence 5971, Ap
					Sequence 12914, A
					Sequence 12570, A
					Sequence 2, Appli
					Sequence 524, Appli
					Sequence 28, Appl
					Sequence 793, Appl
					Sequence 3, Appli
					Sequence 8080, Ap
					Sequence 8235, Ap
					Sequence 23, Appli

C 13	46	20.7	2409	4	US-09-230-225B-3	Sequence 3, Appli
14	46	20.7	3489	2	US-08-728-323A-1	Sequence 1, Appli
15	46	20.7	3489	4	US-09-298-568-1	Sequence 1, Appli
16	46	20.7	3489	4	US-09-410-399-1	Sequence 1, Appli
C 17	46	20.7	4897	6	5196516-7	Patent No. 5196516
18	46	20.7	5207	2	US-09-858-664A-1	Sequence 1, Appli
19	46	20.7	6638	2	US-08-070-301-2	Sequence 2, Appli
C 20	46	20.7	32207	2	US-08-770-379-20	Sequence 20, Appli
C 21	46	20.7	32207	3	US-08-757-669A-20	Sequence 20, Appli
C 22	46	20.7	32207	4	US-09-330-371A-20	Sequence 20, Appli
23	46	20.7	77536	4	US-09-410-551B-1	Sequence 1, Appli
24	46	20.7	4403765	3	US-09-103-840A-2	Sequence 1, Appli
25	46	20.7	4411529	3	US-09-103-840A-1	Sequence 1, Appli
C 26	45	20.3	840	4	US-09-252-991A-2027	Sequence 2027, Ap
C 27	45	20.3	867	2	US-08-961-858-2	Sequence 2, Appli
C 28	45	20.3	867	2	US-08-961-858-4	Sequence 2, Appli
C 29	45	20.3	867	3	US-09-089-593-2	Sequence 2, Appli
C 30	45	20.3	867	3	US-09-089-593-4	Sequence 4, Appli
C 31	45	20.3	867	3	US-08-950-925-3	Sequence 3, Appli
C 32	45	20.3	867	4	US-09-565-286-3	Sequence 3, Appli
33	45	20.3	1065	4	US-09-252-991A-1930	Sequence 1930, Ap
34	45	20.3	1125	4	US-09-252-991A-11557	Sequence 11557, A
C 35	45	20.3	1203	4	US-09-252-991A-9574	Sequence 9574, Ap
C 36	45	20.3	1404	4	US-09-252-991A-4420	Sequence 4420, Ap
C 37	45	20.3	1551	4	US-09-252-991A-12931	Sequence 12931, A
C 38	45	20.3	1620	4	US-09-252-991A-16490	Sequence 16490, A
C 39	45	20.3	1659	4	US-09-252-991A-13081	Sequence 13081, A
C 40	45	20.3	1845	4	US-09-614-034-188	Sequence 188, App
C 41	45	20.3	2073	4	US-09-252-991A-12858	Sequence 12858, A
C 42	45	20.3	2133	4	US-09-252-991A-16218	Sequence 16218, A
43	45	20.3	2346	4	US-09-252-991A-9636	Sequence 9636, Ap
44	45	20.3	2358	4	US-09-252-991A-16381	Sequence 16381, A
45	45	20.3	2454	4	US-09-252-991A-4186	Sequence 4186, Ap

ALIGNMENTS

RESULT 1
US-09-313-294A-5971
; Sequence 5971, Application US/09313294A

; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7690
; SOFTWARE: PERL Program
; SEQ ID NO 5971
; LENGTH: 284
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700350992H1
; NAME/KEY: unsure
; LOCATION: 70, 83, 239, 258-259, 283
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-5971

Alignment Scores:	0.000278	Length:	284
Pred. No.:	48.00	Matches:	38
Score:	69.57%	Conservative:	10
Best Local Similarity:	55.07%	Mismatches:	21
Query Match:	21.62%	Indels:	0
DB:	4	Gaps:	0

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-313-294A-5971 (1-284)

QY 39 LysileProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAla 58
 Db 3 AAGTCCCCCAAGACCACTGGTATGCTTTTGTGATTTGAGATCTCTCGTGTCT 62
 QY 59 AspAspAlaileTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGlu 78
 Db 63 GAGGAGGNAATTGCTGAACGNGATGGATACAACTTTGATGACACCGCTCTAAGAGTGGAG 122
 QY 79 IleAlaHisGlyGlyArgArg*****SerTyrSer 98
 Db 123 GCTGCTCATGTTGGTGTAGAGTAATGCTTCTCGCATGATGCTTCAAGTGGCTTGGTGGCG 182
 QY 99 AlaSerArgAlaProSerArgArgSer 107
 Db 183 GTGTGGAGCAGTACGTGGTGTGTGCG 209

RESULT 2
 ; Sequence 12914, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 12914
 ; LENGTH: 888
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (111)
 ; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
 US-09-252-991A-12914

Alignment Scores:
 Pred. No.: 0.0887 Length: 888
 Score: 48.00 Matches: 38
 Percent Similarity: 21.92% Conservative: 10
 Best Local Similarity: 17.35% Mismatches: 171
 Query Match: 21.62% Indels: 0
 DB: 4 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-252-991A-12914 (1-888)

QY 3 SerArgTyrAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgGlyCysGlu 22
 Db 198 AGCGGTACACGCGCGTGTGGCGTCGATCCCGACGCTCGACAGGAACAGCCGAGCAGCG 257
 QY 23 ValGluAspLeuPheTyrIleTyrGlyProIleValAlaPheAspLeuIleProPro 42
 Db 258 CCGCAGCAGGAGTCTTCAGCGCGCGTCCCGGCCATCCCGCAGGAGCAGATGGCGA 317
 QY 43 ArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaile 62
 Db 318 ACACCATCAGCAGCAATATCCCGCGGACCGAAGCGCATCCGACCTTGGCCAGCAGCG 377
 QY 63 TyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly 82
 Db 378 GGGCGAAGCAGGACCATCCCGCAGGTGGCGATGAAAGCGCAATCAACGAACCTCCAGGCCG 437
 QY 83 GlyArgArg*****SerTyrSerAlaSerArgAla 102
 Db 438 ACAGCAGACAGGCCCG 497
 QY 103 ProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTyr 122

Db 498 TCATCAGGTCGAGGCTTCGCGGGATGTTGAGCAGGATCGAGTGTATCCCGCGCGCT 557
 QY 123 GluAspLeuLeuAspHisMetArgGlyAlaGlyAspValCysPheSerGluValPhePro 142
 Db 558 ATTTCGAGCCCAAGGTATACGCGCGGAGCAGGATCAGCGCGGTTCGCGCGGCGAGCGGA 617
 QY 143 AspArgGlyGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetIleTyrAla 162
 Db 618 GGGCGAATCCACCGGATGATGCGCGCGGATGCGCGCGGTGATCGACCGAGGCTGCGCAGA 677
 QY 163 IleArgGlyLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgVal 182
 Db 678 GGCTACGATGTCGCGATCAGGTCGCGATCAGCGCGGTGATCAGTGTGGTGGCGGTCA 737
 QY 183 ArgGluTyrGluSerArgSerValSerArgSerProAspAspSerIleSerTyrArgSer 202
 Db 738 TCGCAGCGCGAAGCGCTGCGCGAGTAGTTCAAGGTTTCCATCAGCAAGTCTCTCACAGC 797
 QY 203 ArgSerArgSerArgGlyProSerCysSerTyrSerSerIleSerArgSerValSer 221
 Db 798 GGCGAAGTGCAGAACGCCCGCAGCGGAGCGTCCAGGCGCGGTGCGAACAGC 854

RESULT 3
 ; Sequence 12570, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 12570
 ; LENGTH: 1824
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (976),(1124)
 ; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
 US-09-252-991A-12570

Alignment Scores:
 Pred. No.: 3.38 Length: 1824
 Score: 48.00 Matches: 38
 Percent Similarity: 21.92% Conservative: 10
 Best Local Similarity: 17.35% Mismatches: 171
 Query Match: 21.62% Indels: 0
 DB: 4 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-252-991A-12570 (1-1824)

QY 3 SerArgTyrAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgGlyCysGlu 22
 Db 889 AGCGGTACACGCGCGTGTGGCGTCGATCCCGACGCTCGACAGGAACAGCCGAGCAGCG 830
 QY 23 ValGluAspLeuPheTyrIleTyrGlyProIleValAlaPheAspLeuIleProPro 42
 Db 829 CCGCAGCAGGAGTCTTCAGCGCGCGTCCCGGCCATCCCGCAGGAGCAGATGGCGA 770
 QY 43 ArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaile 62
 Db 769 ACACCATCAGCAGCAATATCCCGCGGACCGAAGCGCATCCGACCTTGGCCAGCAGCG 710
 QY 63 TyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly 82

Db 709 GGGCGAAGACCATCCCGAGTGGCGATGAAGCGCGCAATGAACGAATCCAGCGCG 650
Qy 83 GlyArgArg*****SerTyrSerAlaSerArgAla 102
Db 649 ACAGCAGAGGGCCACCGCGCGAGCCCTTGGTGCCTCGCGTACCGGTGAGCGGTGG 590
Qy 103 ProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrp 122
Db 589 TCATCAGGTGAGGCTTCCCGGGGATGTTGAGCAGGATCGAGCTGATCCCGCCCGCGT 530
Qy 123 GlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPhePro 142
Db 529 ATTCCGAGCCAGGTATACGGCGGAGCAGGATCAGCGGGTTTCCGCGCGAGCGCGA 470
Qy 143 AspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspMetLysTyrAla 162
Db 469 GGGCGAATGCCACCGGATCAGCAGTCCACCGCGTTGATCGAGCCAGCGCTGCAGCA 410
Qy 163 IleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgVal 182
Db 409 GGCTACGATGGTCCGATCAGGGTCCCATGACCGCGTGCATGAGTTGGTCCGCGTCA 350
Qy 183 ArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSer 202
Db 349 TCGGACGCGAAGCCTCGCCGAGTAGTTGAGGTTTCCATCAGCAAGTCTCTACAGC 290
Qy 203 ArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSer 221
Db 289 GGCAGGAAGTCGAGAAGCCGACGCGAGCGGACGTCGAGCGCGGTCGGAACAGC 233

RESULT 4
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 1.12e+06 Length: 4403765
Score: 48.00 Matches: 38
Percent Similarity: 22.02% Conservative: 10
Best Local Similarity: 17.43% Mismatches: 170
Query Match: 21.62% Indels: 0
DB: 3 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-103-840A-2 (1-4403765)
Qy 3 SerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGlu 22
Db 2435386 TCACCTTCAGCGGACAGCGGCGGTGAGTCGAGTCGACATCCAGCCCGGTAGACGCTCAGG 2435327
Qy 23 ValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProPro 42
Db 2435326 TGCGCGCGCTGGCGATCTTGGCTGGACCGACAGATAGCGGTGTCCACGACGCGCTCTCCA 2435267

Qy 43 ArgProProGlyTyrAlaPheValGluPheValGluPheValGluPheValGluPheValGlu 62
Db 2435266 ACGGTCCTCTCGCTCGCAGCTTCGACACGCGGTTCATCGTAGTGGCGCGCGATTCGAT 2435207
Qy 63 TyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly 82
Db 2435206 GCAGACGGAAGCCTAACGGCGATGCTGCGCGGACACCGAGTCTACACTTGGCGGGT 2435147
Qy 83 GlyArgArg*****SerTyrSerAlaSerArgAla 102
Db 2435146 GGGCAGCATTTCCCGCTCGCATGACGTTTTTGGACCTTGACCAACCGACGACGACCTGCC 2435087
Qy 103 ProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrp 122
Db 2435086 CCGGTCGCGGAGTCTCGCGGTTCGCGTCCAGCAAGTGGCCAGCAGTCCGCGGAGG 2435027
Qy 123 GlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPhePro 142
Db 2435026 TCATCTGGTGGCGGTGGCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2434967
Qy 143 AspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspMetLysTyrAla 162
Db 2434966 CTCGGTCAGGTGTCAGAGCCTGCGCGGTGTTGACCATCTGTCAGCAGCGGGGCTA 2434907
Qy 163 IleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgVal 182
Db 2434906 TCGGACACCGAGATTATGCGTGGCTTTTACCCCGATCCGTCGCTGACCATCACCCG 2434847
Qy 183 ArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSer 202
Db 2434846 CGCGGTCGCGTGCACCGCTGACCAACCGCTCCGTCGTCGATGATGATGATGATGATGAT 2434787
Qy 203 ArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220
Db 2434786 GGCTCGCGAGTGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 2434733

RESULT 5
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Alignment Scores:
Pred. No.: 1.12e+06 Length: 4411529
Score: 48.00 Matches: 38
Percent Similarity: 22.02% Conservative: 10
Best Local Similarity: 17.43% Mismatches: 170
Query Match: 21.62% Indels: 0
DB: 3 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-103-840A-1 (1-4411529)
Qy 3 SerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGlu 22
Db 2438086 TCACCTTCAGCGGACAGCGGCGGTGAGTCGAGTCGACATCCAGCCCGGTAGACGCTCAGG 2438027

```

; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 524
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(866)
US-09-620-312D-524

Alignment Scores:
Pred. No.: 0.298      Length: 955
Score: 47.00         Matches: 37
Percent Similarity: 22.17%      Conservative: 10
Best Local Similarity: 17.45%    Mismatches: 165
Query Match: 21.17%             Indels: 0
DB: 4                       Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-620-312D-524 (1-955)

QY 23 ValGluAspLeuPheTyrLysTy-GlyProIleValAspIleAspLeuLysIleProPro 42
DB 2438026 TGCCECCGCTCGCATCTTGCCTCGACCATAGCGCTGTCCAGCAGCGCGCTCTCCA 2437967
QY 43 ArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIle 62
DB 2437966 ACGGTCCTCGTGCAGCTTCGACCGGGTCTATGTAGTGGCGCGCGCATGGAT 2437907
QY 63 TyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly 82
DB 2437906 GCAGAACGGAAGCCTAACCGCATCGCTGGCGCGCACCGAGCTCTACACTTGGCGGGT 2437847
QY 83 GlyArgArg*****SerTyrSerAlaSerArgAla 102
DB 2437846 GCGCAGCATTCGCGTGCATGACGTTTGGACCCCTGACGACCGACGACGACCTGCC 2437787
QY 103 ProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrp 122
DB 2437786 CCGGTCGCCGAGCTGCTCGCGGTTCCGGTCAGCAAGGTGGCCAGCAGCTGGGGGAGG 2437727
QY 123 GlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPhePro 142
DB 2437726 TCATCTGGTGGCGCGCGCGCTGGTGGCGGTGGTATCCCTCAGCTCTCTTCACCAA 2437667
QY 143 AspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAla 162
DB 2437666 CTCGGTCAGGTGTCAGAGCCTCGCGGCTGTTCACCATCTGCACGACGGGGCTA 2437607
QY 163 IleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgVal 182
DB 2437606 TCGCGCACCCAGAGATTATCGCTGGCTGTTTACCCCGCATCCGTCGTCACCATC 2437547
QY 183 ArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSer 202
DB 2437546 CGAGCGTCGCGTGACGCGCGTCAGCAACGCCCGTCGCGTCGATGCATGATGCCACCA 2437487
QY 203 ArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220
DB 2437486 GGCTCGCAGGTGGTGGCGCGCGCCAGGCTATGGCGTATGGCACCAGCGGGT 2437433
```

RESULT 6

```

US-09-620-312D-524
; Sequence 524, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR FILING DATE: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
```

```

; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 524
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(866)
US-09-620-312D-524

Alignment Scores:
Pred. No.: 0.298      Length: 955
Score: 47.00         Matches: 37
Percent Similarity: 22.17%      Conservative: 10
Best Local Similarity: 17.45%    Mismatches: 165
Query Match: 21.17%             Indels: 0
DB: 4                       Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-620-312D-524 (1-955)

QY 11 ValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyrLysTyr 30
DB 129 GTGCGGAGCGCCACCGAGCTTGAGGACTTGGCCGCTGAGTTGGTTCGATATGGCCCTATA 188
QY 31 GlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAlaPheVal 50
DB 189 GTAGACGTTTACATTCCTCCTTACACTCTCCCGCCCAAGAGGATTTGCTTATGTT 248
QY 51 GluPheGluAspProArgAspAlaIleTyrGlyArgAspGlyTyrAspPhe 70
DB 249 CAATTGAGATGTTCCGATGCTGAAGATGCTCTTTTATACCTCAATAGAAAGTGGGTA 308
QY 71 AspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArgArg***** 90
DB 309 TGTGCGCTCAGATTGAAATCAGTTTCACAAAGGTGATCGCAAAACACCGAGCCCAATG 368
QY 91 *****SerTyrSerAlaSerArgAlaProSerArgSerArgSerArg 110
DB 369 AAATCAAAAGAACGTCATCTTCTCAAGTATGATCAGGAGATCAAGAGAGCCCGCAGC 428
QY 111 ValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArg 130
DB 429 CAAAGAAGAACTCGAAGTAGAAGTTCTTCATGGGGAAGAAATAGAGGGGTTCAGACAGC 488
QY 131 LysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyVal 150
DB 489 CTTAAAGAGTCTCGACACAGGGGATTTCTTATAGCCAGTCTAAATCTCGTTCCTCAATCA 548
QY 151 ValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGlu 170
DB 549 TTACCAAGCGCGTCTACCTCAGCAAGGCGAGTCAAGAACTCCCAAGAGGAATTTTGGCTCT 608
QY 171 PheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArgSerVal 190
DB 609 AGAGGACGCTCAAGGTCCAAGTCTTACAAAAGAGGTCCAAAGTCAATAGGAAATACACAG 668
QY 191 SerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGlyProSer 210
DB 669 TCAAGTTCACCTCAAAAGCAGACTAGCTCAGGAAACAAATCAAGATCATCATGGAAGACAT 728
QY 211 CysSerTyrSerSerLysSerArgSerValSerPro 222
DB 729 TCTGACTCAATAGCAAGATCCCGGTGTAATCTCC 764
```

RESULT 7

```

US-09-180-109A-28
; Sequence 28, Application US/09180109A
; Patent No. 6410293
; GENERAL INFORMATION:
; APPLICANT: MUKUMOTO, Fujio
; APPLICANT: NISHIO, Shoichi
; APPLICANT: AKIMARU, Jiro
; APPLICANT: MITSUDA, Satoshi
```

; TITLE OF INVENTION: DNA Fragments Containing Biotin Biosynthetase Gene and
; FILE REFERENCE: Use of the Same
; CURRENT APPLICATION NUMBER: US/09/180,109A
; PRIORITY FILING DATE: 1998-12-03
; PRIORITY FILING DATE: 1997-03-03
; PRIORITY FILING DATE: 1997-03-03
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 28
; LENGTH: 1336
; TYPE: DNA
; ORGANISM: Spingomonas paucimobilis
; FEATURE:
; OTHER INFORMATION: Strain = JCM7511
; NAME/KEY: CDS
; LOCATION: (151)..(1173)
US-09-180-109A-28

Alignment Scores:
Pred. No.: 1.63 Length: 1336
Score: 47.00 Matches: 37
Percent Similarity: 21.86% Conservative: 10
Best Local Similarity: 17.21% Mismatches: 168
Query Match: 21.17% Indels: 0
DB: Gaps: 4

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-180-109A-28 (1-1336)

QY 6 AsnArgThrIleTyrValGluAsnLeuProGlyAspIleArgLysCysGluValGluAsp 25
DB 6 AACAGGACTATCAGGCACTTACGATGCGCGGTACAGCGATTTTCGGTCCCGCACCA 65
QY 26 LeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProProArgPro 45
DB 66 ATCTTGTAAGCGCGGAGGATGCTAAGCTGCTGGGACATATATGCGCCCGAGG 125
QY 46 GlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArg 65
DB 126 CGGGCGAATCAGAGGACGACATGATGACGACGACACCGCGCTGAGTCCGAGCGAC 185
QY 66 AspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyArgArg 85
DB 186 CCGCGCACCGACTGACCGCGCGGAGATCGCGGCTGTTTCGACCTCCCTTCCACGA 245
QY 86 *****SeryTyrSerAlaSerArgAlaProSerArg 105
DB 246 GCTGTTGTTCCGCGCGCGGAGTGACCGCGCATCACCGCGCGATCAGGTTCACT 305
QY 106 ArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeu 125
DB 306 GTCCAGCTGTTGTCATCAAGACGGCGGCTGCCCGAGGATTGGGCTATTGACGCA 365
QY 126 LysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLys 145
DB 366 GTCCAGCCATCGCATACCGGCTGAAGCGACCAAGCTGATGGACCGCGCGCTGCT 425
QY 146 GlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLys 165
DB 426 GCAGGCGGCGCGCAGCGCAAGGATCAGCGCTCGACCGCTTCTGCTGGGCGCGCTG 485
QY 166 LeuAspAlaThrCluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyr 185
DB 486 GGCACACCCCAAGGATCGCGACATGCCCGCATCGTGGAGATGGTGAAGGGCGTCCGCG 545
QY 186 GluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArg 205
DB 546 CATGGGCATGGAACCTGATGACGCTGGGATGTCGACCGATGACAGGCGCACACGCT 605
QY 206 SerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220
DB 606 CGCCGAGGCGGCGCTGACTATTACAAATCAATATCGACACGTC 650

RESULT 8
US-09-252-991A-793/c
; Sequence 793, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 793
; LENGTH: 2508
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-793

Alignment Scores:
Pred. No.: 39.4 Length: 2508
Score: 47.00 Matches: 37
Percent Similarity: 21.76% Conservative: 10
Best Local Similarity: 17.13% Mismatches: 169
Query Match: 21.17% Indels: 0
DB: Gaps: 4

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-252-991A-793 (1-2508)

QY 7 ArgThrIleTyrValGluAsnLeuProGlyAspIleArgLysCysGluValGluAspLeu 26
DB 1322 CGCAGGCACTCGGTTCTCGAAGCTGAGCGGCCCTCTCGTATTGCGGCGACGACGCAA 1263
QY 27 PheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProProArgProGly 46
DB 1362 AGATTCAACAGCGGTCATGTCAGCGCGGCTCGAAGTCCGCGACGCGCGGTCCACG 1203
QY 47 TyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAsp 66
DB 1202 AGTTGGCGGTGGCATGCGCGAGCATCTGTTGCGCGCGCGGTTGGCGTAGCGTACGCGG 1143
QY 67 GlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyArgArg*** 86
DB 1142 CTGTCCAGTTGACCCAGAGGATGCCAGGGTGTGTTGTCGATGGAGAACTGGGTCCAGG 1083
QY 87 *****SeryTyrSerAlaSerArgAlaProSerArgArg 106
DB 1082 CGCAGCGCTCTCTCGCGCGGCTGCGGAGTCCAGGTCCGCGCGGCGCGGTACAGCGCG 1023
QY 107 SerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLys 126
DB 1022 GTTTCAGCGCATGCTGCTGCGCGGCTGGAGAAACACCGCAGCGCGCGCGCGCGCAGC 963
QY 127 AspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGly 146
DB 962 AGCAGGCGGAGGAGGATCGACAGGTTCTGCCAGAACCCCGCGCGACTCGCGAGCGCGGA 903
QY 147 MetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeu 166
DB 902 TACTTGGTTGCAACACGCGCTCGTCAACTGGTCCAGGAGCTTGGCGGACGAGGAGCGC 843
QY 167 AspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGlu 186
DB 842 AACGCTGGTGCACCACTCGCAGTTCCGGCCAGTCCGCGACGGGTGCCACCGCGTAGC 783
QY 187 SerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSer 206
DB 782 AGCTGCGGTAGCGGATATCGCAACACCGCTAGCTCGCGCAACTCCGTTCCGCGGTC 723
QY 207 ArgGlyProSerCysSerTyrSerSerLysSerArgSerValSerPro 222

Db 203 ACCAGCACACGTCGCGCTCCACGGCGGCATCCAGCAGCGCTGGTGGCAGTG 144
Qy 183 ArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSer 202
Db 143 ACCCGCTGACGATCGCGCGCACCTCGCGCGCGCTTCGACCTGCAGGCGGTCGGGCAG 84
Qy 203 ArgSerArgSerArgGlyProSerCysSerTyrSerLysSerArgSerVal 220
Db 83 TAGTCCTGGATGCGCGCGGCATCCAGATAACGGTCGGCCTCTTCGACACGGGTG 30
RESULT 11
US-09-252-991A-8235
; Sequence 8235, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 8235
; LENGTH: 888
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8235
Alignment Scores:
Pred. No.: 0.481 Length: 888
Score: 46.00 Matches: 36
Percent Similarity: 21.10% Conservative: 10
Best Local Similarity: 16.51% Mismatches: 172
Query Match: 20.72% Indels: 0
DB: 4 Gaps: 0
US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-252-991A-8235 (1-888)
Qy 3 SerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGlu 22
Db 176 TCGGTGCGATGATGTCGGCGGGGATGAGCTGATGCGCTTCTCCGGCGCGTGTGACG 235
Qy 23 ValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProPro 42
Db 236 GTCTGCTCGAACCCTCGCGGTCAGGTAGGCATCGACGCCGCGCATCGCTGCTG 295
Qy 43 ArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIle 62
Db 296 ATATAGCCTTGGCGCGCGGTCACCGAGGATGCGCGGATGCGCTGACCGGCATCG 355
Qy 63 TyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly 82
Db 356 ACCAGCAGCGGTCACGCCCGCCGCGATCGGCACATCGCGCGGAAATCGCGAGCTGC 415
Qy 83 GlyArgArg*****SerTyrSerAlaSerArgAla 102
Db 416 ATCGGCTCGCAAGCAACGAGGAGGACGATGGAACGGGGTTGCCGGCTCCAAAGCGC 475
Qy 103 ProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrp 122
Db 476 CCTTCCACTTCGAAGCGGAGCTCGCGGGGAGTGTCGACCTTGTGGCCACCTCCGGATGC 535
Qy 123 GlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPhePro 142
Db 536 AGATCAGCGGCGAGTGTGAAGCGGAGCAGACTGATATCGTTGTTCAGAGGCTCTTCAG 595
Qy 143 AspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAla 162

Db 596 CGCGTGTGCTTCATCGCGCACCGCAGGGGTTCACACCTCCAGAAAGTAACCGTGATGC 655
Qy 163 IleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgVal 182
Db 656 ACCAGCACACGTCGCGCTCCAGCGCGCATCCAGCAGCGCTGGTGGCAGTG 715
Qy 183 ArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSer 202
Db 716 ACCCGCTGAGGATCGCGCGGCACCTCGCGCGGCTTCGACCTGCAGCGCGTTCGGGCAG 775
Qy 203 ArgSerArgSerArgGlyProSerCysSerTyrSerLysSerArgSerVal 220
Db 776 TAGTCCTGGATGCGCGCGGCATCCAGATAACGGTCGGCCTCTTCGACACGGGTG 829
RESULT 12
US-09-180-109A-23
; Sequence 23, Application US/09180109A
; Patent No. 6410293
; GENERAL INFORMATION:
; APPLICANT: NUKUMOTO, Fujio
; APPLICANT: NISHIO, Shoichi
; APPLICANT: AKIMARU, Jiro
; APPLICANT: MITSUDA, Satoshi
; TITLE OF INVENTION: DNA Fragments Containing Biotin Biosynthetase Gene and
; TITLE OF INVENTION: Use of the Same
; FILE REFERENCE: 0152-0490P
; CURRENT APPLICATION NUMBER: US/09/180,109A
; CURRENT FILING DATE: 1998-12-03
; PRIOR APPLICATION NUMBER: 09/047838 JAPAN
; PRIOR FILING DATE: 1997-03-03
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 23
; LENGTH: 1420
; TYPE: DNA
; ORGANISM: Sphingomonas paucimobilis
; FEATURE:
; OTHER INFORMATION: Strain = JCM7511
; NAME/KEY: CDS
; LOCATION: (223)..(1245)
US-09-180-109A-23
Alignment Scores:
Pred. No.: 5.17 Length: 1420
Score: 46.00 Matches: 36
Percent Similarity: 21.50% Conservative: 10
Best Local Similarity: 16.82% Mismatches: 168
Query Match: 20.72% Indels: 0
DB: 4 Gaps: 0
US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-180-109A-23 (1-1420)
Qy 7 ArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeu 26
Db 81 AGGACTATCAGGCATCTACGATCGCGGGTACAGCGGATTTTCGGTCCCGGCACCAATC 140
Qy 27 PheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProProArgProGly 46
Db 141 TTGTAAAGCGCGCGAGGATGCTGAGGTGCTGGGACATATATATCCGCCCGGAGCGG 200
Qy 47 TyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAsp 66
Db 201 GCGAATGACAGGACACACGCTGATGACGACGACCGCGCTGAGCTCCGAGGCGACCCC 260
Qy 67 GlyTyrAspPheAspGlyCysArgLeuArgValIleAlaHisGlyArgArg*** 86
Db 261 GCGCACCAGCTGGAGCGCGCGGATGCGCGGCTGTTCCGACCTGCCCTTCCAGGCT 320
Qy 87 *****SerTyrSerAlaSerArgAlaProSerArgArg 106
Db 321 GTTCTTCGCGCGCGAGGTGACCGCGCGCATCACCGCGCGATCAGGTTAGCTGTC 380
Qy 107 SerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLys 126

Db 381 GACGCTGTTCGATCAAGAGCGGGCGGTGCCCGAGGATTGGCGTATTGCACCCAGTC 440
QY 127 AspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGly 146
Db 441 GACCCATGCCATACCGGCTCAAGCGACCAAGTGTATGACCGCGCGGTGTGCA 500
QY 147 MetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeu 166
Db 501 GCGCGCGCGCAGGCAAGATCAAGTCTCAGCGCGCTTTCGATGGCGCGCGCTGGCG 560
QY 167 AspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGlu 186
Db 561 CAACCCCAAGATCCGACATCCCGCATCGTGGAGATGTTGAAGGCGTGGCGGCAT 620
QY 187 SerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSer 206
Db 621 GGGCATGAAACCTGCATGACGCTGGCATGCTGACCGATGCACAGGCGCAGCGCTCGC 680
QY 207 ArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220
Db 681 CGAGCGGGGTGGACTATTACATCAATATCGACACGTC 722

RESULT 13
US-09-230-225B-3/c
; Sequence 3, Application US/09230225B
; Patent No. 6403362
; GENERAL INFORMATION:
; APPLICANT: Meiji Seika Kaisha, Ltd.
; APPLICANT: Moriya, Tatsuki
; TITLE OF INVENTION: Systems for the Mass Production of Proteins or Peptides by Micro
; TITLE OF INVENTION: of the Genus Humicola
; FILE REFERENCE: VX990054
; CURRENT APPLICATION NUMBER: US/09/230,225B
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 2409
; TYPE: DNA
; ORGANISM: Humicola insolens
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (389)..(457)
; OTHER INFORMATION:
; NAME/KEY: mat_peptide
; LOCATION: (458)..()
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (458)..(477)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (478)..(535)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (536)..(1029)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (1030)..(1141)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (1142)..(1761)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (1762)..(1815)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (1816)..(1989)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (1990)..(2044)
; OTHER INFORMATION:
; NAME/KEY: CDS

; LOCATION: (2045)..(2095)
; OTHER INFORMATION:
; NAME/KEY: misc_feature
; LOCATION: (688)..(693)
; OTHER INFORMATION: Cleavage site SmaI
; NAME/KEY: misc_feature
; LOCATION: (1253)..(1259)
; OTHER INFORMATION: Cleavage site BamHI
; NAME/KEY: misc_feature
; LOCATION: (1505)..(1510)
; OTHER INFORMATION: Cleavage site BglII
; NAME/KEY: misc_feature
; LOCATION: (1643)..(1648)
; OTHER INFORMATION: Cleavage siteStuI
US-09-230-225B-3
Alignment Scores:
Pred. No.: 74.9 Length: 2409
Score: 46.00 Matches: 36
Percent Similarity: 21.80% Conservative: 10
Best Local Similarity: 17.06% Mismatches: 165
Query Match: 20.72% Indels: 0
DB: 4 Gaps: 0
US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-230-225B-3 (1-2409)
QY 11 ValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyrLysTyr 30
Db 1230 GTTGGCGCGTGTGTTGGGATCGCCACTCGCGCTTCAAGCGCGGCGCAGCGACGTCGCG 1171
QY 31 GlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAlaPheVal 50
Db 1170 GTCAGGAAGTCTGTAACGACGATCGGCTGGATTTGGCGTTAATTTGCGTCCAGAATG 1111
QY 51 GluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyrAspPhe 70
Db 1110 ATGTGGAGATGTGTGGTATTTGATGATGTGTGGTGCAGGGGTACAGGTTAGGTTGG 1051
QY 71 AspGlyCysArgLeuArgValGluIleAlaHisGlyArgArg***** 90
Db 1050 TGGTGTGACACCGCACITACCGGCATACGGGGGTTCGGGCGCCCTGGTTTCGGCGCGC 991
QY 91 *****SerTyrSerAlaSerArgAlaProSerArgSerArgSerArg 110
Db 990 GGATCTCAGAGAGGCTCTCGACGAGCGGTGTCCACCGTGCACGTTGGCGTGCAGCCACT 931
QY 111 ValLeuValThrGlyLeuProProSerAlaSerTyrGlnAspLeuLysAspHisMetArg 130
Db 930 GGAAGCTCGGACCTCAGCGACGCGCGGCGGCGCCCTCAAGCGAGGTCGGTGTGATCT 871
QY 131 LysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyVal 150
Db 870 GAGGAATGCGAGGGTGTGGACCTCAGAGCGGTAGTAGTTGTTGGCCAGAGCTCGACAC 811
QY 151 ValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGlu 170
Db 810 CCTCGAAGGGTTCGGTGTGTAGTGGCGGTGTGGTGGCCACCGCAGGATGTGTGACGG 751
QY 171 PheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArgSerVal 190
Db 750 TCGGGGTGGGAGCAGTGTGTGATCGAGGTCAACACCGGTCCTGTGTGTGCGAGTGGGCC 691
QY 191 SerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGlyProSer 210
Db 690 GGGAGGTGTGTGACGAGCTCGAAGTGCAGCTAGTGTGAGGTCTGTGTGTGCTGTGTGTG 631
QY 211 CysSerTyrSerSerLysSerArgSerValSer 221
Db 630 GCAAGCAGTGGGAGTACCAGTCTGTTCTGTTCA 598
RESULT 14
US-08-728-323A-1
; Sequence 1, Application US/08728323A


```

Db 1208 AGGAGGATGACGATGATGAGGACAAATGAGGACGAGGAGGATGACGAGGAGGAGGACAAAGA 1267
      |||
Qy 84 ArgArg*****SerTyrSerAlaSerArgAlaPro 103
      |||
Db 1268 AGGAGGACGAGGAGGACGGGGCGGATGGAAACAAACGTTGAGCATCCAAAGTTCACAAC 1327
      |||
Qy 104 SerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGln 123
      |||
Db 1328 AGCAGCAGGAGCCCAACACAGCAGGAGCCACAGCAGCAGGAGGCCACACAGCAGGAGGCCCC 1387
      |||
Qy 124 AspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAsp 143
      |||
Db 1388 TGCAGGAGCCCAACACAGCAGGAGCCACAGCAGGAGGCCACACAGCAGGAGGCCCTGC 1447
      |||
Qy 144 ArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIle 163
      |||
Db 1448 AGGAGCCCAACACAGCAGGAGCCACAGCAGCAGGAGGCCCTGCAGGAGGCCCAACACAGCAGG 1507
      |||
Qy 164 ArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArg 183
      |||
Db 1508 AGCCACAACAGCAGGAGGCCACAGCAGCAGGAGGCCACAGCAGCAGGAGGCCACACAGCAGG 1567
      |||
Qy 184 GluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArg 203
      |||
Db 1568 AGCCACAGCAGCAGGAGGCCACAGCAGCAGGAGGCCACAGCAGCAGGAGGCCACACAGCAGG 1627
      |||
Qy 204 SerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSer 221
      |||
Db 1628 AGCCACAGCAGCAGGAGGCCACAGCAGCAGGAGGCCACAGCAGCAGGAGGCCACACAGCAGG 1681
      |||

```

Search completed: February 4, 2004, 21:53:57
 Job time : 1275 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

QM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2004, 20:08:19 ; Search time 266 Seconds
(without alignments)
2252.914 Million cell updates/sec

Title: US-10-014-927-19MOD_COPY_1_222

Perfect score: 222

Sequence: 1 MSSRWNRRTYVGNLPDGRK.....RSRSGPSCSVSSKSRVSP 222

Scoring table: UNITARY2

Xgapop 10.0 , Xgapext 0.5

Xgapop 10.0 , Xgapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2_1/USPTO_spool/BAUM927/runat_04022004_131333_2700/app_query.fasta_1.391
-DB=N_Geneseq_19Jun03 -QFWT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=unitary2 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=BAUM927 @CGN 1.1.312 @runat_04022004_131333_2700 -NCFU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	222	100.0	771	21	AAC49242	Arabidopsis thalia
2	222	100.0	925	21	AAC49248	Arabidopsis thalia
3	222	100.0	1213	21	AAC40180	Arabidopsis thalia
4	218	98.2	714	21	AAC42865	Arabidopsis thalia
5	207	93.2	762	24	AB213394	Arabidopsis thalia
6	143	64.4	1232	21	AAC34814	Arabidopsis thalia
7	141	63.5	1190	21	AAC38528	Arabidopsis thalia
8	130.5	58.8	1241	21	AAC37026	Arabidopsis thalia
9	116.5	52.5	885	21	AAC42654	Arabidopsis thalia
10	98	44.1	479	21	AAC35763	Zea mays DNA fragm
11	95	42.8	1428	22	AAC402904	Human shear stress
12	90	40.5	1458	21	AAC44527	Zea mays DNA fragm
13	87.5	39.4	1069	24	ABN96845	Human prostate can
14	87.5	39.4	1162	21	AAF16226	Human prostate can
15	87.5	39.4	1218	23	ABV30279	Human prostate exp
16	87.5	39.4	1475	24	ABO54393	Human ovarian anti
17	87.5	39.4	1584	24	ABY93368	Mouse ischaemic co
18	81	36.5	345	24	ABQ85913	Arabidopsis thalia
19	73.5	33.1	446	24	ABY94679	Human pancreatic c
20	73.5	33.1	1478	23	ABL29431	Drosophila melanog
21	73	32.9	492	21	AAC47771	Zea mays DNA fragm
22	69	31.1	759	24	AB577425	Frog embryonic gen
23	69	31.1	2601	24	AB235222	Human gene express
24	65	29.3	742	24	AB577306	Frog embryonic gen
25	63	28.4	495	24	ABY99907	Mouse ischaemic co
26	63	28.4	497	21	AAC09125	Human secreted pro
27	62	27.9	2076	24	ABK83839	Human cDNA differe
28	61.5	27.7	444	23	ABV45807	Human prostate exp
29	60.5	27.3	727	23	ABV16009	Human prostate exp
30	56	25.2	4044	21	AAC81899	A. thaliana SRP30
31	55.5	25.0	1454	24	ABK84288	Human cDNA differe
32	55.5	25.0	1534	25	ABX63268	Human cDNA #268 di
33	55.5	25.0	3697	25	ABX62342	Human activated T
34	55.5	25.0	3948	23	ABL29430	Drosophila melanog
35	55	24.8	1644	23	ABL30487	Drosophila melanog
36	55	24.8	2367	20	AAZ41956	Human myometrium t
37	54	24.3	471	21	AAC41641	Arabidopsis thalia
38	54	24.3	1150	21	AAC51497	Arabidopsis thalia
39	54	24.3	1153	21	AAC33304	Arabidopsis thalia
40	53	23.9	568	21	AAC34694	Arabidopsis thalia
41	53	23.9	913	20	AAC33485	Human prostate can
42	53	23.9	57248	24	ABK83563	Human cDNA differe
43	52	23.4	913	21	AAC34439	Arabidopsis thalia
44	52	23.4	1365	21	AAC51484	Arabidopsis thalia
45	52	23.4	1367	21	AAC33948	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAC49242
ID AAC49242 standard; DNA; 771 BP.
AC AAC49242;
XX
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 60449.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
XX EF1033405-A2.

XX	06-SEP-2000.	PR	02-JUL-1999;	99US-0142055.
PD		PR	06-JUL-1999;	99US-0142390.
XX		PR	08-JUL-1999;	99US-0142803.
PF		PR	09-JUL-1999;	99US-0142920.
XX		PR	12-JUL-1999;	99US-0142977.
XX		PR	13-JUL-1999;	99US-0143542.
PR		PR	14-JUL-1999;	99US-0143624.
PR		PR	15-JUL-1999;	99US-0144005.
PR		PR	16-JUL-1999;	99US-0144086.
PR		PR	19-JUL-1999;	99US-0144325.
PR		PR	19-JUL-1999;	99US-0144331.
PR		PR	19-JUL-1999;	99US-0144332.
PR		PR	19-JUL-1999;	99US-0144333.
PR		PR	19-JUL-1999;	99US-0144334.
PR		PR	19-JUL-1999;	99US-0144335.
PR		PR	20-JUL-1999;	99US-0144352.
PR		PR	20-JUL-1999;	99US-0144632.
PR		PR	20-JUL-1999;	99US-0144634.
PR		PR	20-JUL-1999;	99US-0144884.
PR		PR	21-JUL-1999;	99US-0144814.
PR		PR	21-JUL-1999;	99US-0145086.
PR		PR	21-JUL-1999;	99US-0145088.
PR		PR	22-JUL-1999;	99US-0145085.
PR		PR	22-JUL-1999;	99US-0145087.
PR		PR	22-JUL-1999;	99US-0145089.
PR		PR	22-JUL-1999;	99US-0145192.
PR		PR	23-JUL-1999;	99US-0145145.
PR		PR	23-JUL-1999;	99US-0145218.
PR		PR	23-JUL-1999;	99US-0145224.
PR		PR	26-JUL-1999;	99US-0145276.
PR		PR	27-JUL-1999;	99US-0145913.
PR		PR	27-JUL-1999;	99US-0145918.
PR		PR	27-JUL-1999;	99US-0145919.
PR		PR	28-JUL-1999;	99US-0145951.
PR		PR	02-AUG-1999;	99US-0146386.
PR		PR	02-AUG-1999;	99US-0146388.
PR		PR	02-AUG-1999;	99US-0146389.
PR		PR	03-AUG-1999;	99US-0147038.
PR		PR	04-AUG-1999;	99US-0147204.
PR		PR	04-AUG-1999;	99US-0147302.
PR		PR	05-AUG-1999;	99US-0147192.
PR		PR	05-AUG-1999;	99US-0147260.
PR		PR	06-AUG-1999;	99US-0147303.
PR		PR	06-AUG-1999;	99US-0147416.
PR		PR	09-AUG-1999;	99US-0147493.
PR		PR	09-AUG-1999;	99US-0147935.
PR		PR	10-AUG-1999;	99US-0148171.
PR		PR	11-AUG-1999;	99US-0148319.
PR		PR	12-AUG-1999;	99US-0148341.
PR		PR	13-AUG-1999;	99US-0148565.
PR		PR	13-AUG-1999;	99US-0148684.
PR		PR	16-AUG-1999;	99US-0149368.
PR		PR	17-AUG-1999;	99US-0149175.
PR		PR	18-AUG-1999;	99US-0149426.
PR		PR	20-AUG-1999;	99US-0149722.
PR		PR	20-AUG-1999;	99US-0149723.
PR		PR	20-AUG-1999;	99US-0149929.
PR		PR	23-AUG-1999;	99US-0149902.
PR		PR	23-AUG-1999;	99US-0149930.
PR		PR	25-AUG-1999;	99US-0150566.
PR		PR	26-AUG-1999;	99US-0150884.
PR		PR	27-AUG-1999;	99US-0151065.
PR		PR	27-AUG-1999;	99US-0151066.
PR		PR	27-AUG-1999;	99US-0151080.
PR		PR	30-AUG-1999;	99US-0151303.
PR		PR	31-AUG-1999;	99US-0151438.
PR		PR	01-SEP-1999;	99US-0151930.
PR		PR	07-SEP-1999;	99US-0152363.
PR		PR	10-SEP-1999;	99US-0153070.
PR		PR	13-SEP-1999;	99US-0153758.
PR		PR	15-SEP-1999;	99US-0154018.
PR		PR	16-SEP-1999;	99US-0154039.
PR		PR	20-SEP-1999;	99US-0154779.

```

PR 22-SEP-1999; 99US-01551139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 1-26e-42 Length: 771
Score: 222.00 Matches: 212
Percent Similarity: 100.00% Conservative: 10
Best Local Similarity: 95.50% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x AAC49242 (1-771)
QY 1 MetSerArgTriaAsnArgThrIleTyrValGlyValAsnLeuProGlyAspIleArgLys 20
DB 1 ATGAGTAGCGGATGGATCGTACGATCTACGTGGGAATTGCTGGAGATATTCGCGAG 60
QY 21 CysGluValGluAspLeuPheTyrIysTyrGlyProIleValAspIleAspLeuLysIle 40
DB 61 TGTGAGGTTGAAGATCTCTTACAGATGATGACCAATTGGCAATTGGAATTGAAGATT 120
QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
DB 121 CCACCGAGACCTCTGTTATGCCCTTTCGAGTTGAAGATCCCTCGTGATGCAGACCAT 180
QY 61 AlaIleTyrGlyValArgAspGlyTyrAspPheAspGlyCysAtgLeuArgValGluIleAla 80
DB 181 GCAATTATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 81 HisGlyGlyValArgArg*****SerTyrSerAlaSer 100
DB 241 CATGTGTGTCGTAGATTTCACCATCAGTTGATAGTACAGCAGCAGCAGTACAGTGCAGC 300
QY 101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla 120
DB 301 CGTGACACCTTCAAGACCGCTCTGACTACCGCTGCTGTGACCGGATTACCGCTTCTGCT 360

```

```

QY 121 SerTyrGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
DB 361 TCGTGGCAGGACCTTAAGGATCATCGCGCAAGCTGGAGATGCTGCTTCTCTGAAGTT 420
QY 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
DB 421 TTCCTGACCGTAAAGGCATGTCGGGGTTGGATTATAGCAACTATGATGATATGAAG 480
QY 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
DB 481 TACGCAATAAGGAACTTGATGCCACTGAATTCGAAATCCTTCTCTAGTCTTATATA 540
QY 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
DB 541 CGGTGAGGAATATGATGTCAGAGTGTGAGTCGAGCCAGCCAGATGATTTCTAAAGCTAT 600
QY 201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220
DB 601 AGAAGCAGGAGTCGGAGCCGCTGGTCCCAAGCTGTAGCTATATAGTACCAAGACGAGGTGTG 660
QY 221 SerPro 222
DB 661 TCACCT 666

RESULT 2
AAC49248
ID AAC49248 standard; DNA; 925 BP.
AC AAC49248;
XX
XX 18-OCT-2000 (first entry)
DT Arabidopsis thaliana DNA fragment SEQ ID NO: 60471.
DE Arabidopsis thaliana
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.

```


PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135363.
PR 24-MAY-1999; 99US-0135628.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140699.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142058.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144362.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152263.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.

```
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 2,398-42 Length: 925
Score: 222.00 Matches: 212
Percent Similarity: 100.00% Conservative: 10
Best Local Similarity: 95.50% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x AAC49248 (1-925)

QY 1 MetSerArgTyrAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
Db 134 ATGAGTAGCGGATGAATCGTACGATCTACGTTGGGAATTTGCCTGGAGATATTCGCAAG 193
QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40
Db 194 TGTGAGGTGAAGATCTCTTCTACAGTATGACCAATTTGGACATTTGATTTGAAGATT 253
QY 41 ProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
Db 254 CCACCGAGACCTCCTGGTATGCCCTTTCGAGTTTGAGATCCTCGTGATCGACGACAT 313
QY 61 AlalleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
Db 314 GCAATTTATGACGCTGATGTTATGATTTGATGGTGTGACATTCGGGTTGAGATTGCA 373
QY 81 HisGlyGlyArgArg+*****+*****+*****+*****+SerTyrSerAlaSer 100
Db 374 CATGTGGTTCGAGATTTTCCACATCAGTTGATAGGTACAGCAGCAGCTACAGTCCGAGC 433
QY 101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla 120
Db 434 CGTGACCTTTCAGACGCTCTGACTACCGCGTGTGTGACCGGATTTACCGCCCTTCTGCT 493
QY 121 SerTyrGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
Db 494 TCGTGCAGGACCTTAAGATCAGATCGCAAGCTGGAGATGCTGCTTCTCTGAAGTT 553
QY 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
Db 554 TTCCCTGACCGTAAAGGATGCTGGGTTGTGGATTATAGCAACTATGATGATATGAAG 613
QY 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
Db 614 TACGCAATAAGAAACTTGAATCCACTGAATTTGAAATGCTTTCTAGTGCCTATATA 673
QY 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
Db 674 CGGGTAGGGAATATAGTCGAGGAGTGTGATCGAAGCCAGATGATTTCAAAGCTAT 733
QY 201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220
Db 734 AGAAGCAGAGTCGGAGCGCTGCTCCAGCTGTAGCTATAGTAGCAAGCAGGAGTGTG 793
QY 221 SerPro 222
Db 794 TCACCT 799

RESULT 3
AAC40180
ID AAC40180 standard; DNA; 1213 BP.
XX
AC AAC40180;
XX
```

```
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 27352.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; terminator sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
```

PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140931.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	16-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144684.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160814.
PR	23-JUL-1999;	99US-0145145.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160981.
PR	26-JUL-1999;	99US-0145276.	PR	22-OCT-1999;	99US-0160989.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145919.	PR	25-OCT-1999;	99US-0161406.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161359.
PR	02-AUG-1999;	99US-0146385.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146388.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161920.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161992.
PR	04-AUG-1999;	99US-0147302.	PR	28-OCT-1999;	99US-0161993.
PR	05-AUG-1999;	99US-0147192.	PR	29-OCT-1999;	99US-0162142.
PR	05-AUG-1999;	99US-0147260.	Alignment Scores:		
PR	06-AUG-1999;	99US-0147303.	Pred. No.:		
PR	06-AUG-1999;	99US-0147416.	Score:		
PR	09-AUG-1999;	99US-0147493.	Percent Similarity:		
PR	09-AUG-1999;	99US-0147935.	Best Local Similarity:		
PR	10-AUG-1999;	99US-0148171.	Query Match:		
PR	11-AUG-1999;	99US-0148319.	DB:		
PR	12-AUG-1999;	99US-0148341.	6.18e-42		
PR	13-AUG-1999;	99US-0148565.	222.00		
PR	16-AUG-1999;	99US-0149368.	100.00%		
PR	17-AUG-1999;	99US-0149375.	95.50%		
PR	18-AUG-1999;	99US-0149426.	100.00%		
PR	20-AUG-1999;	99US-0149722.	21		
PR	20-AUG-1999;	99US-0149723.	Length:		
PR	20-AUG-1999;	99US-0149929.	Matches:		
PR	23-AUG-1999;	99US-0149902.	Conservative:		
PR	23-AUG-1999;	99US-0149930.	Mismatch:		
PR	26-AUG-1999;	99US-0150568.	Indels:		
PR	26-AUG-1999;	99US-0150884.	Gaps:		
PR	27-AUG-1999;	99US-0151065.	US-10-014-927-19MOD_COPY_1_222 (1-222) x AAC40180 (1-1213)		
			QY		
			1 MetSerSerArgTipAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20		
			DB		
			134 ATGAGTAGCCGATGAATCGTACGATCTCTTACAGTATGACCAATTGGACATTGTTGAAGATT 193		
			QY		
			21 CysGluValGluAspLeuPheTyrIleValAspIleAspLeuLysIle 40		
			DB		
			194 TGTGAGTTGAAGATCTCTTACAGTATGACCAATTGGACATTGTTGAAGATT 253		
			QY		
			41 ProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60		
			DB		
			254 CCACCGAGACCTCTCGTATTGCTTTCGAGTTTGAAGATCCTCGTATGACAGCAT 313		

```
Qy 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
Db |||||
314 GCATTTATGACGTGATGATTGATTTGATGGGTGCGATTCGGTTGAGATTGCA 373
Qy 81 HisGlyGlyArgArg*****SerTyrSerAlaSer 100
Db |||||
374 CATGGTGGTGTAGATTTTCCACCATCAGTTGATAGGTACAGCAGCAGCTACAGTGGGAGC 433
Qy 101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla 120
Db |||||
434 CGTGCACCTTCAAGACCTCTGACTATCGCGTCTTGTGACCGGATTTACCGCCTTCTGCT 493
Qy 121 SerTyrGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
Db |||||
494 TCGTGGCAGGACCTTAAGGATCATCGCGAAGCTGGAGATGTCGCTTCTGAGGT 553
Qy 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
Db |||||
554 TTCCTGACCGTAAGGACGTCTGGGTGTGATTAAGCAACTATGATGATGATGAG 613
Qy 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
Db |||||
614 TACGCAATAAGAACTTGTATGACCACTGAAATTCGAATGCTTCTCTAGTGTCTATATA 673
Qy 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
Db |||||
674 CGGTGAGGGAATATGAGTCTGAGGAGTGTGAGTGAAGCCAGATGATTTCTAAAGCTAT 733
Qy 201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220
Db |||||
734 AGAAGCAGGAGTGGAGCCGTGTCCAGCTGTAGCTATAGTACCAAGACGAGGAGTG 793
Qy 221 SerPro 222
Db |||||
794 TCACCT 799

RESULT 4
AAC42865
ID AAC42865 standard; DNA; 714 BP.
XX
AC AAC42865;
XX
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 37144.
XX
DE
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 23-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
```

```
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145114.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 9.64e-42 Length: 714
Score: 218.00 Matches: 208
Percent Similarity: 100.00% Conservative: 10
Best Local Similarity: 95.41% Mismatches: 0
Query Match: 98.20% Indels: 0
DB: 21 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x AAC42865 (1-714)

QY 1 MetSerSerArgTTPAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
DB 1 ATGAGTAGCCGATGGAATCGTAGCATCTACGTGGAAATTCCTCGGAGATATTCGCAAG 60
QY 21 CysGluValGluAspLeuPheTyrIleTyrGlyProIleValAspIleAspLeuLysIle 40
DB 61 TGTGAGTTGAGATCTCTTCTACAGATATGACCAATTTGGACATTTGATTTGAGATT 120
QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
DB 121 CCACCGAGACCTCTCGTTATGCTTCGAGTTTGAAGATCCTCGTGATCGACAGCAT 180
QY 61 AlalleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
DB 181 GCAATTTATGACGCTGATGTTATGATTTGATGGGTGTCGACTTCGGGTTGAGATTGCA 240
QY 81 HisGlyGlyArgArg******SerTyrSerAlaSer 100
DB 241 CATGTTGTTGTTAGATTTTCACCATCAGTTAGGTACAGCAGCAGCTACAGTCGAGC 300
QY 101 ArgAlaProSerArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla 120
DB 301 CGTGACCTTCAGACGCTCTGACTACCGCGTGTGTGACCGGATTAACCGCTTCGCT 360
QY 121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
DB 361 TCGTGGCAGGACCTTAGGATCACATGGCGAAGCTGGAGATGCTGCTTCTCTGAAGTT 420
QY 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
DB 421 TTCCCTGACCGTAAAGGATCTCTGGGTTGTGATTTATACCACTATGATGATGATGAAG 480
QY 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
DB 481 TACGCAATAGGAACCTTGATGCGCACTGAAATTTGAAATGCTTCTCTAGTGTCTATATA 540
QY 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
DB 541 CGGGTGAGGGAATATGATGCTCGAGGAGTGTGAGTCGAGCCCGCATGATTTCTAAAAGCTAT 600
QY 201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArg 218
DB 601 AGAGCAGGAGTCGAGCGCGTGGTCCCAAGCTGTAGCTATAGTAGCAGAGCAGG 654
```

```

RESULT 5
ABZ13394 standard; DNA; 762 BP.
XX
AC ABZ13394;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1199.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
FN W0200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26685.
XX
PR 24-AUG-2000; 2000US-227866P.
PR 26-JAN-2001; 2001US-264647P.
PR 22-JUN-2001; 2001US-300111P.
XX
PA (SCRI ) SCRIPPS RES INST.
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Krops J, Wang X, Zhu T;
XX
DR WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed
PT and producing plants with increased tolerance to these abiotic stresses
PT
XX
PS Claim 144; SEQ ID NO 1199; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising:
CC (a) contacting nucleic acid representative of expressed polynucleotides
CC in the plant cell with an array of probes representative of the plant
CC cell genome; and
CC (b) detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention.
CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.
XX
SQ Sequence 762 BP; 189 A; 161 C; 197 G; 215 T; 0 other;

Alignment Scores:
Pred. No.: 6.86e-39 Length: 762
Score: 207.00 Matches: 197
Percent Similarity: 100.00% Conservative: 10
Best Local Similarity: 95.17% Mismatches: 0
Query Match: 93.24% Indels: 0
DB: 24 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x ABZ13394 (1-762)
QY 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
DB 1 ATGAGTAGCGATGGAATCGTACGATCAGCTACGTTGGGAATTCCTGGAGATTCGCAAG 60
QY 21 CysGluValGluAspLeuPheTyrIleTyrGlyProIleValAspIleAspLeuLysIle 40
DB 61 TGTGAGGTGAAGATCTCTTCTACAAAGTATGACCAATTTGGACATTGGAATTGAAGATT 120

```

```

QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
DB 121 CCACCGAGACCTCTCTGTTATGCTTTTGTGAGTTTGAAGATCTCTGTTGAGCAGCAT 180
QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
DB 181 GCAATTTATGGACGTGATGCTTATGATTTTGTGAGGTGTCGACTTCGGTTGAGATTGCA 240
QY 81 HisGlyIlyArgArg*****SerTyrSerAlaSer 100
DB 241 CATGGTGTGTAGATTTTCCACATCAGTTGATAGGTACAGCAGCAGCTACAGTGGCAGC 300
QY 101 ArgAlaProSerArgArgSerAspTyrAcqValIleValThrGlyLeuProProSerAla 120
DB 301 CGTGCACTTCAAGACGCTCTGACTACCGCGTCTTGTGACCGGATTACCGCCTTCTGCT 360
QY 121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
DB 361 TCGTGGCAGGACCTTAAGGATCAGATGCGCAAGCTGGAGATGCTCTCTGAAGTT 420
QY 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
DB 421 TTCCTGACCGTAAAGGCATGCTCTGGGTTGTGGATTATAGCAACTATGATGATATGAAG 480
QY 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
DB 481 TACGCAATAAGSAACTTGATGCCACTGAAATTCGAAATGCTTCTCTAGTCTTATATA 540
QY 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
DB 541 CGGTTGAGGAAATATGATGTCAGGAGTGTGATGTCGAGAGCCAGATGATTTCTAAAGCTAT 600
QY 201 ArgSerArgSerArgSerArg 207
DB 601 AGAAGCAGGAGTCGAGCCGT 621

RESULT 6
AAC34814
ID AAC34814 standard; DNA; 1232 BP.
XX
AC AAC34814;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 7984.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
FN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.

```

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139432.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Alignment Scores:			
Pred. No.:	3, 87e-22	Length:	1232
Score:	143.00	Matches:	148
Percent Similarity:	68.40%	Conservative:	10
Best Local Similarity:	64.07%	Mismatches:	63
Query Match:	64.41%	Indels:	10
DB:	21	Gaps:	1
US-10-014-927-19MOD_COPY_1_222 (1-222) x AAC34814 (1-1232)			
QY	1	MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys	20
DB	94	ATGAGCAGTCGTCGAGTAGACCGGTGACGTCGGAACCTCTCGCGATATCCGTGAG	153
QY	21	CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle	40
DB	154	AGAGAGTCCAGATTGTTTCAGTAGTAGACCTGTGTTTCAAATTGATTGAAGGTT	213
QY	41	ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp	60
DB	214	CTCCCAAGGCTCTCGGTATGATTCGTTGATTCGATGATGCTCGGATGCTGAAGAT	273
QY	61	AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla	80
DB	274	GCTATTATGCTGATGCTATGATTTGATGGGATCGTTTGAAGGCTGGAATTGGCG	333
QY	81	HisGlyGlyArgArg-----*****	90
DB	334	CATGGTGGAGGCGTTTCATCAGATGATCTCGGGTAGTTTCAATGGTGGCGCGTGT	393
QY	91	*****SerTyrSerAlaSerArgAlaProSerArgArgSerAspTyrArg	110
DB	394	GGTGGTGGCGCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	453
QY	111	ValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArg	130
DB	454	GTCTAGTCACAGCTTCCTTCATCTGCTTCTTGGCAAGATCTCAAGATCACATGCT	513
QY	131	LysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyVal	150
DB	514	AAAGGAGCGCATGCTGTTCTCGCAAGTGTACCGTGATGCTAGAGGACCACTGGAGTT	573
QY	151	ValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGlu	170
DB	574	GTGTATTACACCTGCTATGAGGACATGATGTCGCTGAAAGCTGACGACACAGAG	633
QY	171	PheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArgSerVal	190
DB	634	TTTCGAATCGTTTCGATGATGATGTCGCGGTAGAGATATGATTCAGGAAGGAT	693
QY	191	SerArgSerProAspAspSerLysSerTyrArgSerArgSerArgGlyProSer	210
DB	694	TCTAGGAGTCTAGCGCGGGAAGATCTATTCTAAGAGCGCGAGCGCGCGTGGACGA	753

QY	211	CysSerTyrSerSerLysSerArgSerValSer	221
DB	754	ACCGTGAGCCGAGCAGGAGCAGAGCAGGAGC	786
RESULT 7			
AAAC38528			
ID	AAC38528	standard; DNA; 1190 BP.	
XX	XX		
AC	AAC38528;		
XX	XX		
DT	17-OCT-2000	(first entry)	
XX	XX		
DE	Arabidopsis thaliana	DNA fragment SEQ ID NO: 21293.	
XX	XX		
KW	Hybridisation assay; Genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway;		
KW	metabolic pathway; promoter; termination sequence; ss.		
XX	XX		
OS	Arabidopsis thaliana.		
XX	XX		
PN	EP1033405-A2.		
PD	06-SEP-2000.		
XX	XX		
PF	25-FEB-2000; 200EP-0301439.		
XX	XX		
PR	25-FEB-1999;	99US-0121825.	
PR	05-MAR-1999;	99US-0123180.	
PR	09-MAR-1999;	99US-0123548.	
PR	23-MAR-1999;	99US-0125788.	
PR	25-MAR-1999;	99US-0126264.	
PR	29-MAR-1999;	99US-0126785.	
PR	01-APR-1999;	99US-0127462.	
PR	06-APR-1999;	99US-0128234.	
PR	08-APR-1999;	99US-0128714.	
PR	16-APR-1999;	99US-0129845.	
PR	19-APR-1999;	99US-0130077.	
PR	21-APR-1999;	99US-0130449.	
PR	23-APR-1999;	99US-0130510.	
PR	28-APR-1999;	99US-0130891.	
PR	30-APR-1999;	99US-0131449.	
PR	30-APR-1999;	99US-0132048.	
PR	30-APR-1999;	99US-0132407.	
PR	04-MAY-1999;	99US-0132484.	
PR	05-MAY-1999;	99US-0132485.	
PR	06-MAY-1999;	99US-0132486.	
PR	06-MAY-1999;	99US-0132487.	
PR	07-MAY-1999;	99US-0132863.	
PR	11-MAY-1999;	99US-0134256.	
PR	14-MAY-1999;	99US-0134218.	
PR	14-MAY-1999;	99US-0134219.	
PR	14-MAY-1999;	99US-0134221.	
PR	14-MAY-1999;	99US-0134370.	
PR	18-MAY-1999;	99US-0134768.	
PR	19-MAY-1999;	99US-0134941.	
PR	20-MAY-1999;	99US-0135124.	
PR	21-MAY-1999;	99US-0135353.	
PR	24-MAY-1999;	99US-0135629.	
PR	25-MAY-1999;	99US-0136021.	
PR	27-MAY-1999;	99US-0136392.	
PR	28-MAY-1999;	99US-0136782.	
PR	01-JUN-1999;	99US-0137222.	
PR	03-JUN-1999;	99US-0137528.	
PR	04-JUN-1999;	99US-0137502.	
PR	07-JUN-1999;	99US-0137724.	
PR	08-JUN-1999;	99US-0138094.	
PR	10-JUN-1999;	99US-0138540.	
PR	10-JUN-1999;	99US-0138847.	
PR	14-JUN-1999;	99US-0139119.	
PR	16-JUN-1999;	99US-0139452.	
PR	16-JUN-1999;	99US-0139453.	
PR	17-JUN-1999;	99US-0139492.	

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139470.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145092.
PR 23-JUL-1999; 99US-0145132.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145219.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 1.08e-21
Score: 141.00
Length: 1190
Matches: 148
Conservative: 10
Percent Similarity: 67.23%
Best Local Similarity: 62.98%
Mismatch: 63
Query Match: 63.51%
Indels: 14
Gaps: 1
DB: 21

US-10-014-927-19MOD_COPY_1_222 (1-222) x AAC38528 (1-1190)

QY 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20

Db	59	ATGAGCAGTCTGTCGAGTAGACCGTGTACGTCGGAACCTTCTCGCGATATCCGTGAG	118	PR	25-MAR-1999;	99US-0136264.
				PR	29-MAR-1999;	99US-0136785.
Qy	21	CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspLeuLysIle	40	PR	01-APR-1999;	99US-0127462.
				PR	06-APR-1999;	99US-0128234.
Db	119	AGAGAGTCGAAGATTGTTTCAGTAAGTATGACCTGTTGTTCAATGATTGAAGTT	178	PR	08-APR-1999;	99US-0128714.
				PR	16-APR-1999;	99US-0129845.
Qy	41	ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp	60	PR	19-APR-1999;	99US-0130077.
				PR	21-APR-1999;	99US-0130449.
Db	179	CCTCCAAAGGCTCTGCTTATGCATTCGTTGAGTTTATGATGCTCGGGATGCTGAAGAT	238	PR	23-APR-1999;	99US-0130510.
				PR	28-APR-1999;	99US-0130891.
Qy	61	AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla	80	PR	30-APR-1999;	99US-0131449.
				PR	30-APR-1999;	99US-0132048.
Db	239	GCTATTCATGCTGCTGATGCTATGACTTTGATGGGCATGCTTTGAGGGTGAATTGGCG	298	PR	04-MAY-1999;	99US-0132407.
				PR	04-MAY-1999;	99US-0132484.
Qy	81	HisGlyGlyArg-----	86	PR	05-MAY-1999;	99US-0132485.
				PR	06-MAY-1999;	99US-0132486.
Db	299	CATGCTGGAGGCGTTCATCAGATGATACCTCGGGTAGTTTCAATGCTGCTGCGCTGCT	358	PR	07-MAY-1999;	99US-0132487.
				PR	11-MAY-1999;	99US-0132863.
Qy	87	*****SerTyrSerAlaSerArgAlaProSerArgArg	106	PR	14-MAY-1999;	99US-0134218.
				PR	14-MAY-1999;	99US-0134219.
Db	359	GGTGTGCGCGCGCTGCTGATGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	418	PR	14-MAY-1999;	99US-0134221.
				PR	14-MAY-1999;	99US-0134370.
Qy	107	SerAspTyrArgValIleuValThrGlyLeuProSerAlaSerTyrGlnAspLeuLys	126	PR	18-MAY-1999;	99US-0134768.
				PR	19-MAY-1999;	99US-0134941.
Db	419	TCAGAGTTTCGTGCTCTAGTCACAGGCTTCCTTCATCTCTCTCTCTCTCTCTCTCT	478	PR	20-MAY-1999;	99US-0135124.
				PR	21-MAY-1999;	99US-0135353.
Qy	127	AspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGly	146	PR	24-MAY-1999;	99US-0135629.
				PR	25-MAY-1999;	99US-0136021.
Db	479	GATCACAATGTTAAGAGAGCGATGCTGCTTTCGCAAGTGATACCGTATGCTAGAGG	538	PR	27-MAY-1999;	99US-0136392.
				PR	28-MAY-1999;	99US-0136782.
Qy	147	MetSerGlyValValAspTyrSerAsnTyrAspMetLysTyrAlaIleArgLysLeu	166	PR	01-JUN-1999;	99US-0137222.
				PR	03-JUN-1999;	99US-0137528.
Db	539	ACAACTGGAGTGTGTTGATTACACCTGCTATGAGACATGAAGTATGCGCTGAAAGCTC	598	PR	04-JUN-1999;	99US-0137502.
				PR	07-JUN-1999;	99US-0137724.
Qy	167	AspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGlu	186	PR	08-JUN-1999;	99US-0138094.
				PR	10-JUN-1999;	99US-0138540.
Db	599	GACGACACAGAGTTTCGAAATGCGTTTTCGAATGATATGTCGGGTAGAGAATATGAT	658	PR	10-JUN-1999;	99US-0138847.
				PR	14-JUN-1999;	99US-0139119.
Qy	187	SerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSer	206	PR	16-JUN-1999;	99US-0139452.
				PR	16-JUN-1999;	99US-0139453.
Db	659	TCAAGGAAGATTTCTAGAGTCTCTAGCGGGGAGATCCTTATCTTAAGAGCCGACGCCG	718	PR	17-JUN-1999;	99US-0139492.
				PR	18-JUN-1999;	99US-0139454.
Qy	207	ArgGlyProSerCysSerTyrSerSerLysSerArgSerValSer	221	PR	18-JUN-1999;	99US-0139455.
				PR	18-JUN-1999;	99US-0139456.
Db	719	AGCCGTGCGAGACGCTGACCCGAGCAGGACGAGCAGGACGAGCAGGACGAGCAGC	763	PR	18-JUN-1999;	99US-0139457.
				PR	18-JUN-1999;	99US-0139458.
XX	RESULT 8			PR	18-JUN-1999;	99US-0139459.
XX	AAC37026			PR	18-JUN-1999;	99US-0139460.
XX	AAC37026 standard; DNA; 1241 BP.			PR	18-JUN-1999;	99US-0139461.
XX				PR	18-JUN-1999;	99US-0139462.
XX	AC AAC37026;			PR	18-JUN-1999;	99US-0139463.
XX				PR	18-JUN-1999;	99US-0139464.
XX	17-OCT-2000 (first entry)			PR	18-JUN-1999;	99US-0139465.
XX				PR	18-JUN-1999;	99US-0139466.
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 15916.			PR	18-JUN-1999;	99US-0139467.
XX				PR	18-JUN-1999;	99US-0139468.
XX	Hybridisation assay; genetic mapping; gene expression control;			PR	18-JUN-1999;	99US-0139469.
XX	protein identification; signal transduction pathway;			PR	18-JUN-1999;	99US-0139470.
XX	metabolic pathway; promoter; termination sequence; ss.			PR	18-JUN-1999;	99US-0139471.
XX				PR	18-JUN-1999;	99US-0139472.
XX	Arabidopsis thaliana.			PR	18-JUN-1999;	99US-0139473.
XX				PR	18-JUN-1999;	99US-0139474.
XX	EP1033405-A2.			PR	18-JUN-1999;	99US-0139475.
XX				PR	18-JUN-1999;	99US-0139476.
XX	06-SEP-2000.			PR	18-JUN-1999;	99US-0139477.
XX				PR	18-JUN-1999;	99US-0139478.
XX	25-FEB-2000; 2000EP-0301439.			PR	18-JUN-1999;	99US-0139479.
XX				PR	18-JUN-1999;	99US-0139480.
XX	25-FEB-1999; 99US-0121825.			PR	18-JUN-1999;	99US-0139481.
XX	05-MAR-1999; 99US-0123180.			PR	18-JUN-1999;	99US-0139482.
XX	09-MAR-1999; 99US-0123548.			PR	18-JUN-1999;	99US-0139483.
XX				PR	18-JUN-1999;	99US-0139484.
XX	23-MAR-1999; 99US-0125788.			PR	18-JUN-1999;	99US-0139485.

PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	13-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159337.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159338.
PR	21-JUL-1999;	99US-0144814.	PR	14-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	18-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161992.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147204.	PR	29-OCT-1999;	99US-0162142.
PR	04-AUG-1999;	99US-0147302.			
PR	05-AUG-1999;	99US-0147192.			
PR	05-AUG-1999;	99US-0147260.			
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149368.			
PR	17-AUG-1999;	99US-0149175.			
PR	18-AUG-1999;	99US-0149246.			
PR	20-AUG-1999;	99US-0149722.			
PR	20-AUG-1999;	99US-0149723.			
PR	20-AUG-1999;	99US-0149929.			
PR	23-AUG-1999;	99US-0149902.			
PR	23-AUG-1999;	99US-0149930.			
PR	25-AUG-1999;	99US-0150566.			
PR	26-AUG-1999;	99US-0150884.			
PR	27-AUG-1999;	99US-0151065.			
PR	27-AUG-1999;	99US-0151065.			
PR	27-AUG-1999;	99US-0151080.			
PR	30-AUG-1999;	99US-0151303.			
PR	31-AUG-1999;	99US-0151438.			
PR	01-SEP-1999;	99US-0151930.			
PR	07-SEP-1999;	99US-0152363.			
PR	10-SEP-1999;	99US-0153070.			
PR	13-SEP-1999;	99US-0153758.			
PR	15-SEP-1999;	99US-0154018.			
PR	16-SEP-1999;	99US-0154039.			
PR	20-SEP-1999;	99US-0154779.			
PR	22-SEP-1999;	99US-0155139.			
PR	23-SEP-1999;	99US-0155486.			
PR	24-SEP-1999;	99US-0155659.			
PR	28-SEP-1999;	99US-0156458.			
PR	29-SEP-1999;	99US-0156596.			
PR	04-OCT-1999;	99US-0157117.			
PR	05-OCT-1999;	99US-0157753.			
PR	06-OCT-1999;	99US-0157865.			
PR	07-OCT-1999;	99US-0158029.			
			Alignment Scores:		
			Pred. No.:	5,33e-19	Length: 1241
			Score:	130.50	Matches: 137
			Percent Similarity:	64.78%	Conservative: 10
			Best Local Similarity:	60.35%	Mismatches: 67
			Query Match:	58.78%	Indels: 13
			DB:	21	Gaps: 1
US-10-014-927-19MOD_COPY_1_222 (1-222) x AAC37026 (1-1241)					
QY	1	MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgIys	20		
Db	141	ATGAGTGGCGGATTTCTCGTCAATCTATGTTGTAATTCGCGGTGACATTAGGGAA	200		
QY	21	CysGluValGluAspLeuPheTyrIlystTyrGlyProIleValAspIleAspLeuIysIle	40		
Db	201	CATGAGATTGAAGATATCTTTTACAAGTATGCGCCCAITGTGATATTGAATTGAAGTT	260		
QY	41	ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp	60		
Db	261	CCACCTCGGCCTCCATGTTATTTGTTGATTGAGTTCGGATTCGGATCTGAAGAT	320		
QY	61	AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla	80		
Db	321	GCCATCAGGGCGGTGATGGCTATTAATTTGGATGCTGTCGCTTGAGGGTTGAGCTTGA	380		
QY	81	HisGlyGlyArg-----	87		
Db	381	CATGGTGGTCGAGGACAGCTTTCAAGTCATCGTCGTGGTGTACGGTGGTGGTGGCAGC	440		
QY	88	*****	87		
Db	441	GGCTATGGTGGGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	500		
QY	108	AspTyrArgValLeuValThrGlyLeuProProSerAlaSerTyrGlnAspLeuIysAsp	127		
Db	501	GAATTCGGAGTTATTGTACGGGGTCCCATCTGTTTCATGGCAAGATTGAGGAT	560		
QY	128	HisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgIysGlyMet	147		
Db	561	CATATGCGGAAAGCTGGTGTGCTTGTGCTGAGGTGACTCGAGACAGTGATCGAACT	620		

148	SerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysIeuAsp	167
621	TATGGTGTGTGCAGCTACACCAATATGATGACATGAAGTATGCAATAAGGAACTTGAT	680
168	AlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGlnSer	187
681	GACACAGAGTTTCAGAAACCCCTGGGCTAGAGGTTTTATCCGGGTTTAAAGAAATATGAAAGC	740
188	ArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArg	207
741	TCCCGATCAAGAAGCAGAAGCCCAAGCAGAAGCAGAAGTCGAAGTCGTATGCCGAGCCGCT	800

RESULT 9

AAC42654

ID	AAC42654 standard; DNA; 885 BP.
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

XX

AC

XX

DT 17-OCT-2000 (first entry)

xx DE Arabidopsis thaliana DNA fragment SEQ ID NO: 36361.

XX Hybridisation assay; genetic mapping; gene expression control;
KW Hybridisation assay; genetic mapping; gene expression control;
XX
XX

protein identification; signal transduction pathway; metabolic pathway; promoter; termination sequence; ss.

XX
OS *Arabidopsis thaliana*.

XX
PN EP1033405-A2.

06-SEP-2000.

Ad XX

XX 25-FEB-2000 200001 0301433
PR 25-FEB-1999 99US-0121825:

05-MAR-1999; 99US-0123180.
09-MAR-1999; 99US-0123548.

PR	23-MAR-1999;	99US-0125788.
PR	25-MAR-1999;	99US-0126264.

PR	23-MAR-1999;	99US-0120204
PR	29-MAR-1999;	99US-0126785
PR	01-APR-1999;	99US-0127462

PR	06-APR-1999;	99US-0128234.
PR	08-APR-1999;	99US-0128714.

00-APR-1999; 99US-0128714
16-APR-1999; 99US-0129845
19-APR-1999; 99US-0130077

EX 13-APR-1999; 99US-0130077
PR 21-APR-1999; 99US-0130449
PR 23-APR-1999; 99US-0130510

PR 23-APR-1999; 99US-0130891.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132407.
PR 30-APR-1999; 99US-0132407.
PR 30-APR-1999; 99US-0132407.
PR 30-APR-1999; 99US-0132408.

FX	30-APR-1999;	99US-0132484.
PR	04-MAY-1999;	99US-0132484.
PR	05-MAY-1999;	99US-0132485.

FK	05-MAY-1999;	99US-0132485
PR	06-MAY-1999;	99US-0132486.
PR	06-MAY-1999.	99US-0132487

PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256

PR II-MAY-1999; 99US-0134238
PR 14-MAY-1999; 99US-0134218
PR 14-MAY-1999; 99US-0134218

PR 14-MAY-1999; 99US-0134219;
PR 14-MAY-1999; 99US-0134221;
PR 14-MAY-1999; 99US-0134370

PR 14-MAY-1999; 99US=0134370;
PR 18-MAY-1999; 99US=0134768;
PR 18-MAY-1999; 99US=0134941

PR	19-MAY-1999;	99US=0134941.
PR	20-MAY-1999;	99US=0135124.
PR	21-MAY-1999;	99US=0135252.

PR 21-MAY-1999; 99US-0135533
PR 24-MAY-1999; 99US-0135629
PR 25-MAY-1999; 99US-0135031

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136793.

PR 28-MAY-1999; 9905-0136/BZ.

PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 5.2e-16 Length: 885
Score: 116.50 Matches: 149
Percent Similarity: 67.66% Conservative: 10
Best Local Similarity: 63.40% Mismatches: 62
Query Match: 52.48% Indels: 15
DB: 21 Gaps: 2

US-10-014-927-19MOD_COPY_1_222 (1-222) x AAC42654 (1-885)

QY 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
DB 1 ATGAGCAGCGGTTTCGAGTAGAAGCATTTACGTCGGGAACCTTCCCGGCGATATCCGTGAA 60
QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40
DB 61 AGAGAAGTTGAAGACTTGTTCAGTAGATGACCTGTGTTTCAATCGATTGAAGATT 120
QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
DB 121 CGCCGAGGCCCTCCAGGCTATGCTTCGCGAGTTTGAGATGCTCGTGATGCTGATGAT 180
QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
DB 181 GCAATTTATGCGCGTGATGTTATGACTTTGATGGCATCATTTACGGGTGGAACCTAGCT 240
QY 81 HisGlyGlyArgArg-----***** 89
DB 241 CATGTTGGAGGCGTTTCATCATGATGATCAGCGGTAGTTATAGTGTGCTGCTGCTGCTGCG 300
QY 90 *****SerTyrSerAlaSerAlaSerAlaProSerArgArgSerArgSerArgTyr 109
DB 301 GGTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360
QY 110 ArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMet 129
DB 361 CGCGTTGTAGTGTACAGTTTGTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 130 ArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGly 149
DB 421 CGTAAAGGAGGAGAGAGTTGTTTCTCAAGTGTTCGTGATGCTGATGATGATGATGATGATG 480
QY 150 ValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIle-----ArgLysLeu 166
DB 481 ATTGTAGATTATACCAGCTACGAGGACATGAAATATGCGGT-GAGATATAAAAAAGCTC 539
QY 167 AspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGlu 186
DB 540 GATGACACAGAGTTTCGGAATCGTTTTCTCATGAATATGTTCCGGTTAGAGATATGAT 599
QY 187 SerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSer 206
DB 600 TCAAGGAGGAGATTCGAGGAGCCCGCGGTGGAAGATCCTATTCTAAGAGCGCGAGCCGT 659
QY 207 ArgGlyProSerCysSerTyrSerSerLysSerArgSerValSer 221
DB 660 GGACGGAGTCCCGCGGTAGTGTGTCGTCGCGCAGCAGAGCAGGAGC 704

RESULT 10
AAC35763
ID AAC35763 standard; DNA; 479 BP.
XX
AC AAC35763;
XX
DT 17-OCT-2000 (first entry)
XX
DE Zea mays DNA fragment SEQ ID NO: 11334.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
OS Zea mays subsp. mays.

XX	EP1033405-A2.	PR	01-JUL-1999;	99US-0141842.
PN		PR	01-JUL-1999;	99US-0142154.
XX		PR	02-JUL-1999;	99US-0142055.
PD	06-SEP-2000.	PR	06-JUL-1999;	99US-0142390.
XX		PR	08-JUL-1999;	99US-0142803.
PF		PR	09-JUL-1999;	99US-0142920.
XX	25-FEB-2000; 2000EP-0301439.	PR	12-JUL-1999;	99US-0142977.
XX		PR	13-JUL-1999;	99US-0143542.
PR	25-FEB-1999;	PR	14-JUL-1999;	99US-0143624.
PR	05-MAR-1999;	PR	15-JUL-1999;	99US-0144005.
PR	09-MAR-1999;	PR	16-JUL-1999;	99US-0144085.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144086.
PR	25-MAR-1999;	PR	19-JUL-1999;	99US-0144325.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144331.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144332.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144333.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144334.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144335.
PR	19-APR-1999;	PR	20-JUL-1999;	99US-0144352.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144884.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144884.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145089.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	PR	03-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	PR	31-AUG-1999;	99US-0151338.
PR	23-JUN-1999;	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	PR	15-SEP-1999;	99US-0154018.

PR	16-SEP-1999;	99US-0154039.	Db	408	CGCGTGGCTCAAAACGTTCTGATTACCGTGTATTGGTTACTGGATTACCTTCTTCAGCA	467
PR	20-SEP-1999;	99US-0154779.				
PR	22-SEP-1999;	99US-0155139.	QY	121	SerTyrGlnAsp	124
PR	23-SEP-1999;	99US-0155486.				
PR	24-SEP-1999;	99US-0155659.	Db	468	TCATGGCAAGAT	479
PR	28-SEP-1999;	99US-0156458.				
PR	29-SEP-1999;	99US-0156596.				
PR	04-OCT-1999;	99US-0157117.	RESULT 11			
PR	05-OCT-1999;	99US-0157753.	AAH02904			
PR	06-OCT-1999;	99US-0157865.	ID	AAH02904	standard; DNA; 1428 BP.	
PR	07-OCT-1999;	99US-0158029.	XX	XX	AAH02904;	
PR	08-OCT-1999;	99US-0158232.	XX	XX	15-JUN-2001 (first entry)	
PR	12-OCT-1999;	99US-0158369.	DT	XX	Human shear stress-response coding sequence SEQ ID NO: 61.	
PR	13-OCT-1999;	99US-0159293.	XX	XX	Human shear stress-response protein; vascular disease;	
PR	13-OCT-1999;	99US-0159294.	DE	DE	arteriosclerosis; ds.	
PR	14-OCT-1999;	99US-0159329.	XX	XX	Human; shear stress-response protein; vascular disease;	
PR	14-OCT-1999;	99US-0159330.	KW	KW	arteriosclerosis; ds.	
PR	14-OCT-1999;	99US-0159331.	XX	XX	Homo sapiens.	
PR	14-OCT-1999;	99US-0159637.	OS	OS	WO200125427-A1.	
PR	14-OCT-1999;	99US-0159638.	PN	PN	12-APR-2001.	
PR	18-OCT-1999;	99US-0159584.	XX	XX	02-OCT-2000; 2000WO-JP06840.	
PR	21-OCT-1999;	99US-0160741.	XX	XX	01-OCT-1999; 99JP-0280976.	
PR	21-OCT-1999;	99US-0160767.	XX	XX	(KYOW) KYOWA HAKKO KOGYO KK.	
PR	21-OCT-1999;	99US-0160768.	XX	XX	(NOJI/) NOJIMA H.	
PR	21-OCT-1999;	99US-0160770.	XX	XX	Nojima H, Yoshiue H, Obayashi M, Ota T, Kawabata A, Sakurada K;	
PR	21-OCT-1999;	99US-0160814.	PI	PI	Kuga T, Sekine S, Nakamura Y, Sugano S;	
PR	21-OCT-1999;	99US-0160815.	XX	XX	WPI; 2001-266308/27.	
PR	22-OCT-1999;	99US-0160880.	DR	DR	P-PSDB; AAH90781.	
PR	22-OCT-1999;	99US-0160981.	XX	XX	DNA sequences, proteins encoded by them and antibodies against them	
PR	22-OCT-1999;	99US-0161404.	PT	PT	useful in diagnosis and treatment of vascular disease caused by	
PR	25-OCT-1999;	99US-0161405.	PT	PT	arteriosclerosis -	
PR	25-OCT-1999;	99US-0161406.	XX	XX	Claim 20; Page 386-388; 678pp; Japanese.	
PR	26-OCT-1999;	99US-0161359.	PS	PS	The present invention provides the protein and coding sequences of a	
PR	26-OCT-1999;	99US-0161360.	XX	XX	number of human shear stress response proteins. These are useful in the	
PR	26-OCT-1999;	99US-0161361.	CC	CC	diagnosis, treatment and screening of vascular diseases caused by	
PR	28-OCT-1999;	99US-0161920.	CC	CC	arteriosclerosis, including heart failure, post-PTCA restenosis and	
PR	28-OCT-1999;	99US-0161992.	CC	CC	hypertension.	
PR	28-OCT-1999;	99US-0161993.	XX	XX	Sequence 1428 BP; 341 A; 296 C; 400 G; 391 T; 0 other;	
PR	29-OCT-1999;	99US-0162142.	SQ	SQ	Alignment Scores:	
					Pred. No.: 6,7e-10 Length: 1428	
					Score: 95.00 Matches: 108	
					Percent Similarity: 56.46% Conservative: 10	
					Best Local Similarity: 51.67% Mismatches: 85	
					Query Match: 42.79% Indels: 6	
					DB: 22 Gaps: 2	
					US-10-014-927-19MOD_COPY_1_222 (1-222) x AAH02904 (1-1428)	
QY	1 MetSerSerArgTyrPAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys	20	QY	9	IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr	28
Db	108 ATGACCCGGCGGAACGGGTGTACGATCTACGTGGCAACCTCCCGCGGACATCCGGAG	167	Db	176	ATCTAGTGGTAACTTACCTCCAGACATCCGACCAAGGACATTGAGACGCTGTCTAC	235
QY	21 CysGluValGluAspLeuPheTyrIleTyrGlyProIleValAspIleAspLeuIle	40	QY	29	LysTyrGlyProIleValAspIleAspLeuIleProProArgProProGlyTyrIle	48
Db	168 AGGGAAGTGGATGATCTCTTCTACAGATGACGCTATAGTGGAAATGACTTGAAAT	227	Db	236	AAATACGGCGCTATCCGCGACATCGACCTCAAGAAATCGCGCGGGGACCGCCCTTCGCC	295
QY	41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp	60	QY	49	PheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyr	68
Db	228 CCACCAAGGCTCCGGTTTGTCTTTTGTGAGTTTGGAGACGACGATGCTGAAGAT	287				
QY	61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla	80				
Db	288 GCAATATATGCCGTGATGATACAACTTTTGATGCCATAGTTTCCGGTGGATTAGCC	347				
QY	81 HisGlyGlyArgArg*****SerTyrSerAlaSer	100				
Db	348 CATGTGACGAGGACATCTCTTTTGTATCGATCTAGCAGCTATAGCAGTGTGTCGACAA	407				
QY	101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProSerAla	120				

Alignment Scores:
Pred. No.: 2.57e-12 Length: 479
Score: 98.00 Matches: 88
Percent Similarity: 79.03% Conservative: 10
Best Local Similarity: 70.97% Mismatches: 26
Query Match: 44.14% Indels: 0
DB: 21 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x AAC35763 (1-479)

Db 296 TTGGTTGAGTTCGAGGACCCGAGACCGCGGTGATGTCGCGACGGCTAT 355
Qy 69 AppPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArg-----Arg 85
Db 356 GATTACGATGGTACCGTCTGCGGTGGAGTTTCCTCGAAGCGGCGGTGGAACAGCCGA 415
Qy 86 *****SerTyrSerAlaSerArgAlaProSerArg 105
Db 416 GCGCGCGCGGGGTGGAGTGGCGAGTCCCGAGGTGCTATGGCCCCCATCCAGG 475
Qy 106 ArgSerAspTyrArgValLeuValThrGlyLeuProPheSerAlaSerTyrGlnAspLeu 125
Db 476 CGGTCTGAACAGAGTGGTGTCTGACTGCCTCCAGTGGAGTGGCAGGATTTA 535
Qy 126 LysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLys 145
Db 536 AAGGATCATCGTGCAGCAGTGTATGTTATGCTGATGTTTACCGAGAT----- 589
Qy 146 GlyMetSerGlyValValAspTyrSerAsnTyrAspMetLysTyrAlaIleArgLys 165
Db 590 ---GGCACTGGTGTCTGGAGTTGTACGGAAAGAGATATGACCTATGCAATTCGAA 646
Qy 166 LeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyr 185
Db 647 CTGGATAACACTAAGTTTAGTCTCATGAGGAGAACTGCCTACATCCGGGTTAAAGTT 706
Qy 186 GluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArg 205
Db 707 GATGGCCCGCAAGTCCAGTTATGGAAGATCTCGATCTCGAAGCGGTAGTGTAGCAGA 766
Qy 206 SerArgGlyProSerCysSerTyrSer 214
Db 767 AGCCGTAGCAGAACCAACAGCAGGAGT 793
RESULT 12
AAC44527
ID AAC44527 standard; DNA; 498 BP.
XX AC AAC44527;
XX DT 18-OCT-2000 (first entry)
DE Zea mays DNA fragment SEQ ID NO: 43156.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
XX
OS Zea mays subsp. mays.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999;	99US-0145085;	PR 21-OCT-1999;	99US-0160768;
PR 22-JUL-1999;	99US-0145087;	PR 21-OCT-1999;	99US-0160770;
PR 22-JUL-1999;	99US-0145089;	PR 21-OCT-1999;	99US-0160814;
PR 22-JUL-1999;	99US-0145192;	PR 21-OCT-1999;	99US-0160815;
PR 23-JUL-1999;	99US-0145145;	PR 22-OCT-1999;	99US-0160980;
PR 23-JUL-1999;	99US-0145218;	PR 22-OCT-1999;	99US-0160981;
PR 23-JUL-1999;	99US-0145224;	PR 22-OCT-1999;	99US-0160989;
PR 26-JUL-1999;	99US-0145276;	PR 25-OCT-1999;	99US-0161404;
PR 27-JUL-1999;	99US-0145913;	PR 25-OCT-1999;	99US-0161405;
PR 27-JUL-1999;	99US-0145918;	PR 25-OCT-1999;	99US-0161406;
PR 27-JUL-1999;	99US-0145919;	PR 26-OCT-1999;	99US-0161359;
PR 28-JUL-1999;	99US-0145951;	PR 26-OCT-1999;	99US-0161360;
PR 02-AUG-1999;	99US-0146386;	PR 26-OCT-1999;	99US-0161361;
PR 02-AUG-1999;	99US-0146388;	PR 28-OCT-1999;	99US-0161920;
PR 02-AUG-1999;	99US-0146389;	PR 28-OCT-1999;	99US-0161920;
PR 03-AUG-1999;	99US-0147038;	PR 28-OCT-1999;	99US-0161992;
PR 04-AUG-1999;	99US-0147204;	PR 28-OCT-1999;	99US-0161993;
PR 04-AUG-1999;	99US-0147302;	PR 29-OCT-1999;	99US-0162142;
PR 05-AUG-1999;	99US-0147302;		
PR 05-AUG-1999;	99US-0147260;	Alignment Scores:	
PR 06-AUG-1999;	99US-0147303;	Pred. No.:	2,96e-10
PR 06-AUG-1999;	99US-0147416;	Score:	90.00
PR 09-AUG-1999;	99US-0147493;	Percent Similarity:	84.56%
PR 09-AUG-1999;	99US-0147935;	Best Local Similarity:	77.21%
PR 10-AUG-1999;	99US-0148171;	Query Match:	40.54%
PR 11-AUG-1999;	99US-0148319;	DB:	21
PR 12-AUG-1999;	99US-0148341;		
PR 13-AUG-1999;	99US-0148565;	US-10-014-927-19MOD_COPY_1_222 (1-222) x AAC44527 (1-498)	
PR 16-AUG-1999;	99US-0148684;	Qy 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys	20
PR 17-AUG-1999;	99US-0149368;	Db 90 ATGAGCAGCGGCAACAGCGGTACCATCTATGTAGCAATCTCCCTGGGACATCCGTGAG	149
PR 18-AUG-1999;	99US-0149426;	Qy 21 CysGluValGluAspLeuPheTyrIleTyrGlyProIleValAspIleAspLeuLysIle	40
PR 20-AUG-1999;	99US-0149722;	Db 150 AGGAGGTGAGGATCTCTTCTACAGTATGCGCGTATTTGGATATTGACTTGAATA	209
PR 20-AUG-1999;	99US-0149723;	Qy 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp	60
PR 23-AUG-1999;	99US-0149902;	Db 210 CCTCGAGACCTCTCGTATAGCAATCGTTGAGTTGAGGATCCACCGTATGCTGATGAT	269
PR 25-AUG-1999;	99US-0150566;	Qy 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluLeuAla	80
PR 26-AUG-1999;	99US-0150884;	Db 270 GCAATTTATGGCGGTGATGGGTATACTTTGATGCTACAGGTTCAGGGTTGAATTAGCT	329
PR 27-AUG-1999;	99US-0151065;	Qy 81 HisGlyGlyArgArg*****SerTy-SerAlaSer	100
PR 27-AUG-1999;	99US-0151066;	Db 330 CATGTGGCAGAGCGCGCTTATTTCTATGATGCTTCAAGCGCTATAGCGTGCATGC	389
PR 27-AUG-1999;	99US-0151080;	Qy 101 Arg---AlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProSer	119
PR 30-AUG-1999;	99US-0151303;	Db 390 CGTGGAGGTGTTCTAGCGCTCTGATTTCCGTGTATGGTCACGTGTTACCTCA-TCG	448
PR 31-AUG-1999;	99US-0151438;	Qy 120 AlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspVal	135
PR 01-SEP-1999;	99US-0151930;	Db 449 GCATCGTGGCAAGATCTGAAGGACCACATGCGCGCGCTGTGTATGTC	496
PR 07-SEP-1999;	99US-0152363;	RESULT 13	
PR 10-SEP-1999;	99US-0153070;	ABN96845	
PR 13-SEP-1999;	99US-0153758;	ID ABN96845 standard; DNA; 1069 BP.	
PR 15-SEP-1999;	99US-0154018;	XX AC ABN96845;	
PR 16-SEP-1999;	99US-0154039;	XX AC ABN96845;	
PR 20-SEP-1999;	99US-0154779;	DT 13-AUG-2002 (first entry)	
PR 22-SEP-1999;	99US-0155133;	XX DE	
PR 23-SEP-1999;	99US-0155486;	XX DE	
PR 24-SEP-1999;	99US-0155659;	XX DE	
PR 28-SEP-1999;	99US-0156458;	XX DE	
PR 29-SEP-1999;	99US-0156596;	XX DE	
PR 04-OCT-1999;	99US-0157117;	XX DE	
PR 05-OCT-1999;	99US-0157153;	XX DE	
PR 06-OCT-1999;	99US-0157865;	XX DE	
PR 07-OCT-1999;	99US-0158029;	XX DE	
PR 08-OCT-1999;	99US-0158232;	XX DE	
PR 12-OCT-1999;	99US-0158369;	XX DE	
PR 13-OCT-1999;	99US-0159293;	XX DE	
PR 13-OCT-1999;	99US-0159294;	XX DE	
PR 13-OCT-1999;	99US-0159295;	XX DE	
PR 14-OCT-1999;	99US-0159329;	XX DE	
PR 14-OCT-1999;	99US-0159330;	XX DE	
PR 14-OCT-1999;	99US-0159331;	XX DE	
PR 14-OCT-1999;	99US-0159637;	XX DE	
PR 14-OCT-1999;	99US-0159638;	XX DE	
PR 18-OCT-1999;	99US-0159584;	XX DE	
PR 21-OCT-1999;	99US-0160741;	XX DE	
PR 21-OCT-1999;	99US-0160767;	XX DE	

Gene #3343 used to diagnose liver cancer.
Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
metastatic liver tumour; cytostatic; expression profile; disease state;
disease progression; drug toxicity; drug efficacy; drug metabolism.

Homo sapiens.
OS
XX
XX
WO200229103-A2.

XX PD 11-APR-2002.
XX PF 02-OCT-2001; 2001WO-US30589.
XX PR 02-OCT-2000; 2000US-237054P.
XX PA (GENE-) GENE LOGIC INC.
XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX DR WPI; 2002-426119/45.
XX PT Diagnosing and detecting the progression of liver cancer,
XX PT hepatocellular carcinoma or metastatic liver tumor in a patient,
XX PT involves detecting the level of expression of two or more genes in a
XX PT liver tissue sample -
XX PS Claim 1; SEQ ID NO 3343; 298pp; English.
XX CC The invention relates to a novel method for diagnosing and detecting the
XX CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX CC tumor in a patient, and differentiating metastatic liver cancer from
XX CC hepatocellular carcinoma in a patient, involving detecting the level of
XX CC expression of two or more genes represented in ABN93503-ABN97455 in a
XX CC tissue sample. The method of the invention has hepatotropic, and
XX CC cytostatic activity. The method is useful for diagnosing and detecting
XX CC the progression of liver cancer, hepatocellular carcinoma and metastatic
XX CC liver carcinoma in a patient. The method is useful for identifying
XX CC expression profiles which serve as useful diagnostic markers as well as
XX CC markers that can be used to monitor disease states, disease progression,
XX CC drug toxicity, drug efficacy and drug metabolism.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 1069 BP; 245 A; 229 C; 310 G; 285 T; 0 other;

Alignment Scores:
Pred. No.: 1-83e-08 Length: 1069
Score: 87.50 Matches: 90
Percent Similarity: 48.54% Conservative: 10
Best Local Similarity: 43.69% Mismatches: 101
Query Match: 39.41% Indels: 5
DB: 24 Gaps: 1

US-10-014-927-19MOD_COPY_1_222 (1-222) x ABN96845 (1-1069)
QY 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
DB 98 AACTAGTGGGAACTTCGACCGACGTGCGGAGAGGACTTGGAGGACCTGTCTTAC 157
QY 29 LysTyrGlyProIleValAspLeuAspLeuLysIleProProArgProGlyTyrAla 48
DB 158 AAGTACGGCGGATCGCGAGATCGAGCTCAAGAACCGGACGGCGCTCGTGGCCCTCGCC 217
QY 49 PheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyr 68
DB 218 TTCGTGCGGTTTCAGGACCCCGAGATCGAGGATGCTATTATGGAAGAATCGTTAT 277
QY 69 AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArgArg***** 88
DB 278 GATATGGCCAGTTCGGCTTCGTGGAG-----TTCCCGAGGACTTAT 322
QY 89 *****SertySerAlaSerArgAlaProSerArgArgSerAsp 108
DB 323 GGAGGTGGGGTGGGGCCCGTGGTGGGAGGAATGGCCCTCTCAAGAAGATCTGAT 382
QY 109 TyrArgValLeuValThrGlyLeuProProSerAlaSerTyrGlnAspLeuLysAspHis 128
DB 383 TTCCGAGTCTTGTTCAGGACTTCCTCGTCAGGACGCTGCGAGGACCTGAAGGATCAC 442
QY 129 MetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSer 149

DB 443 ATGGGAGAGCTGGGATGTCTGTATGCTGATGTCAGAGAGATGGAGTGGGATGGTC 502
QY 149 GlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAla 168
DB 503 GAGTATCTCAGAAAAGAGACATGGAATATGCCCTGCGTAAACTGGATGACACCAATTC 562
QY 169 ThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArg 188
DB 563 CGTCTCTCATGAGGTTGAAACTTCTTATCCGAGTTTATCTCTGAGAGAACACACGCTAT 622
QY 189 SerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGly 208
DB 623 GGCTACTCAGGTTCTGGTCTGGTCAAGGGCGGTGACTCTCATACCAAGACAGGGGT 682
QY 209 ProSerCysSerTyrSer 214
DB 683 TCCCCACACTACTTCTCT 700

RESULT 14
AAFI6226
ID AAFI6226 standard; CDNA; 1162 BP.
XX AC AAFI6226;
XX DT 13-MAR-2001 (first entry)
XX Human prostate cancer antigen nucleotide sequence SEQ ID NO:661.
XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
XX KW neuroprotective; cytostatic; cardiocactive; immunomodulatory; muscular;
XX KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
XX KW antibacterial; gene therapy; neural; immune; reproductive; renal;
XX KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX KW wound; infectious disease; ss.
XX OS Homo sapiens.
XX PN WO200055174-A1.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US05988.
XX PR 12-MAR-1999; 99US-0124270.
XX FA (HUMA-) HUMAN GENOME SCI INC.
XX PA (ROSE/) ROSEN C A.
XX PI Rosen CA, Ruben SM;
XX DR WPI; 2000-587513/55.
XX DR P-PSDB; AAB57023.
XX PT Prostate cancer associated gene sequences, referred to as prostate
XX PT cancer antigens, useful for treatment, prevention, and diagnosis of
XX PT disorders such as prostate cancer -
XX PS Claim 1; Page 1098-1099; 2338pp; English.
XX AAFI5566 to AAFI6505 encode the human prostate cancer associated
XX CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
XX CC The prostate cancer antigens can have neuroprotective, cytostatic,
XX CC cardiocactive, immunomodulatory, muscular, vulnary, gastrointestinal,
XX CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
XX CC and can be used in gene therapy. The prostate cancer antigen
XX CC polynucleotides may be used for detection of prostate cancer, chromosome
XX CC identification, as chromosome markers, and for numerous other diagnostic
XX CC or research purposes. The prostate cancer antigens may be used to treat
XX CC disorders such as neural, immune, muscular, reproductive,
XX CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX CC disorders, wounds, and infectious diseases. AAFI6506 to AAFI6514 to
XX CC AAB57303 represent sequences used in the exemplification of the present

```
CC invention.
XX SQ Sequence 1162 BP; 280 A; 252 C; 340 G; 288 T; 2 other;

Alignment Scores:
Pred. No.: 2,45e-08 Length: 1162
Score: 87.50 Matches: 90
Percent Similarity: 48.54% Conservative: 10
Best Local Similarity: 43.69% Mismatches: 101
Query Match: 39.41% Indels: 5
DB: 21 Gaps: 1

US-10-014-927-19MOD_COPY_1_222 (1-222) x AAF16226 (1-1162)
Qy 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
Db 143 ATCTACGTGGGAACCTTCCGACCGAGTGGCGGAGAGGACTTGGAGGACCTGTCTTCTAC 202
Qy 29 LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAla 48
Db 203 AAGTACGCGCCATCCCGAGATCGAGTCAAGAACCGGACCGCCCTCGGCCCTTCGCC 262
Qy 49 PheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyr 68
Db 263 TTCGTGCGCTTCGAGGACCCCGAGATCGAGATGCTATTATGGAAGAAATGGTTAT 322
Qy 69 AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArgArg***** 88
Db 323 GATTATGCCAGTGTCCGCTTCGTGTGGAG-----TTCGCCAGGACTTAT 367
Qy 89 *****SerTyrSerAlaSerArgAlaProSerArgArgSerAsp 108
Db 368 GGAGTGGGGTGGTGGCCCGTGGTGGGAGGAATGGGCTCTACAGAAGATCTGAT 427
Qy 109 TyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHis 128
Db 428 TTCGAGGTTCTGTTCAGGACTTCCTCCGTCAGGACGCTGGCAGGACCTGAAGAGTAC 487
Qy 129 MetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSer 148
Db 488 ATGGAGAGCTGGGATGCTGTATGCTGATGTGCAGAGATGAGTGGGATGGTC 547
Qy 149 GlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAla 168
Db 548 GAGTATCTCAGAAAAGAGACATGGAATATGCCCTGGCTAACTGGATGACACCAATTC 607
Qy 169 ThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArg 188
Db 608 CGCTCTCATGAGGTGAAACTTCTACATCCGAGTTTATCTCTGAGAGAAGCACCAGCTAT 667
Qy 189 SerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGly 208
Db 668 GGCTACTCACGCTCTCGGTCTGGGTCAAGGGCGCGTGACTCTCCATACCAAGCAGGGT 727
Qy 209 ProSerCysSerTyrSer 214
Db 728 TCCCCACACTACTTCTCT 745

RESULT 15
ID ABV30279 standard; cDNA; 1218 BP.
XX AC ABV30279;
XX DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 30270.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX
```

```
PN W0200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US05171.
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-189862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-219007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX DR WPI; 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX Claim 1; Page 6574-6575; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
XX CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX SQ Sequence 1218 BP; 280 A; 275 C; 357 G; 305 T; 1 other;

Alignment Scores:
Pred. No.: 2,89e-08 Length: 1218
Score: 87.50 Matches: 90
Percent Similarity: 48.54% Conservative: 10
Best Local Similarity: 43.69% Mismatches: 101
Query Match: 39.41% Indels: 5
DB: 23 Gaps: 1

US-10-014-927-19MOD_COPY_1_222 (1-222) x ABV30279 (1-1218)
Qy 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
Db 180 ATCTACGTGGGAACCTTCCGACCGAGTGGCGGAGAGGACTTGGAGGACCTGTCTTCTAC 239
Qy 29 LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAla 48
Db 240 AAGTACGCGCCATCCCGAGATCGAGTCAAGAACCGGACCGCCCTCGTCCCTTCGCC 299
Qy 49 PheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyr 68
Db 300 TTCGTGCGCTTCGAGGACCCCGAGATGCGAGGATGCTATTATGGAAGAAATGGTTAT 359
Qy 69 AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArgArg***** 88
Db 360 GATTATGCCAGTGTCCGCTTCGTGTGGAG-----TTCGCCAGGACTTAT 404
Qy 89 *****SerTyrSerAlaSerArgAlaProSerArgArgSerAsp 108
Db 405 GGAGTTCGGGTGGTGGCCCGCTGGTGGGAGGAATGGGCTCTTCTACAGAAGATCTGAT 464
```

Qy 109 TyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHis 128
Db |||||
465 TTCCGAGTTCTTGTTCAGGACTTCCTCCGTCAGGCAGCTGGCAGGACCTGAAGGATCAC 524
Qy 129 MetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSer 148
Db |||||
525 ATGGGAGAAGCTGGGGATGCTGTATGCTGATGTCAGAAAGGATGGAGTGGGGATGGTC 584
Qy 149 GlyValValAspTyrSerAsnTyrAspMetLysTyrAlaIleArgLysLeuAspAla 168
Db |||||
585 GAGTATCTCAGAAAGAGACATGGAAATATGCCCTGCGTAAACTGGATGACACCAAATTC 644
Qy 169 ThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArg 188
Db |||||
645 CGCTCTCATGAGGGTGAACCTTCATACATCGAGTTTATCCTGAGAGAACCCAGCTAT 704
Qy 189 SerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGly 208
Db |||||
705 GGCTACTCACGGTCTCGGTCTGGGTCAAGGGGCGGTGACTCTCCATACCAAGACAGGGT 764
Qy 209 ProSerCysSerTyrSer 214
Db |||||
765 TCCCACACTACTTCTCT 782

Search completed: February 4, 2004, 21:26:36
Job time : 271 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2004, 20:09:39 ; Search time 349 Seconds
(without alignment)
2343.170 Million cell updates/sec

Title: US-10-014-927-19MOD_COPY_1_222
Perfect score: 222
Sequence: 1 MSSRWNRRTYVGNLPEDIRK.....RSRRGPGSCSYSSKRSVSP 222

Scoring table: UNITARY2
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2449703 seqs, 1841816367 residues 4899406

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/BAUM927/runat_04022004_131334_2737/app_query.fasta_1.391
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPT=0 -UNITS=bits -START=1 -END=1 -MARRX=unitary2
-TRANS=human40.cai -LIST=45 -DOCLAL=LOCAL -OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USRR=BAUM927 @CGN 1 1 221 @runat_04022004_131334_2737
-NCFU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/prodata/1/pubna/US07_PUBCOMB.seq:
2: /cgn2_6/prodata/1/pubna/PCT_NEW_PUB.seq:
3: /cgn2_6/prodata/1/pubna/US06_NEW_PUB.seq:
4: /cgn2_6/prodata/1/pubna/US06_PUBCOMB.seq:
5: /cgn2_6/prodata/1/pubna/US07_NEW_PUB.seq:
6: /cgn2_6/prodata/1/pubna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/prodata/1/pubna/US08_NEW_PUB.seq:
8: /cgn2_6/prodata/1/pubna/US08_PUBCOMB.seq:
9: /cgn2_6/prodata/1/pubna/US09A_PUBCOMB.seq:
10: /cgn2_6/prodata/1/pubna/US09B_PUBCOMB.seq:
11: /cgn2_6/prodata/1/pubna/US09C_PUBCOMB.seq:
12: /cgn2_6/prodata/1/pubna/US09_NEW_PUB.seq:
13: /cgn2_6/prodata/1/pubna/US09A_PUBCOMB.seq:
14: /cgn2_6/prodata/1/pubna/US10A_PUBCOMB.seq:
15: /cgn2_6/prodata/1/pubna/US10B_PUBCOMB.seq:
16: /cgn2_6/prodata/1/pubna/US10_NEW_PUB.seq:
17: /cgn2_6/prodata/1/pubna/US60_NEW_PUB.seq:
18: /cgn2_6/prodata/1/pubna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description

* Query Match Length DB ID Description				

1	207	93.2	762	10	US-09-938-842A-1199	Sequence 1199, Ap
2	207	93.2	762	12	US-09-938-842A-1199	Sequence 1199, Ap
3	95	42.8	3299	13	US-10-006-285-513	Sequence 513, App
4	87.5	39.4	1089	10	US-09-880-107-3342	Sequence 3342, Ap
5	87.5	39.4	1162	10	US-09-925-300-661	Sequence 661, App
6	87.5	39.4	1475	12	US-10-264-049-273	Sequence 273, App
7	87.5	39.4	1717	15	US-10-171-581-9	Sequence 9, Appli
8	81	36.5	345	9	US-09-770-791-783	Sequence 783, App
9	73.5	33.1	446	15	US-10-060-036-52	Sequence 52, Appl
10	69	31.1	759	9	US-09-910-943-651	Sequence 851, App
11	69	31.1	2601	13	US-10-101-510-334	Sequence 334, App
12	65	29.3	742	9	US-09-910-943-539	Sequence 539, App
13	62	27.9	2167	12	US-10-439-703-38	Sequence 38, Appl
14	57	25.7	1579	12	US-10-388-934-197	Sequence 197, App
15	56	25.2	4044	14	US-10-014-927-18	Sequence 18, Appl
16	55.5	25.0	1534	14	US-10-044-090-268	Sequence 268, App
17	55.5	25.0	3697	14	US-10-002-600-58	Sequence 58, Appl
18	53	23.9	3026	11	US-09-919-039-314	Sequence 314, App
19	51	23.0	1750	9	US-09-925-302-316	Sequence 316, App
20	50.5	22.7	489	11	US-09-918-995-23144	Sequence 23144, A
21	50	22.5	9027	13	US-09-930-213-304	Sequence 304, App
22	50	22.5	24387	11	US-09-764-891-9844	Sequence 9844, Ap
23	49	22.1	2212	9	US-09-925-297-142	Sequence 142, App
24	48	21.6	307	9	US-09-294-093B-1314	Sequence 1314, Ap
25	48	21.6	2138	12	US-10-369-493-27664	Sequence 27664, A
26	48	21.6	9400	15	US-10-270-333-7	Sequence 7, Appli
27	48	21.6	9810	11	US-09-373-658-35	Sequence 35, Appl
28	48	21.6	9810	12	US-09-989-687-35	Sequence 35, Appl
29	48	21.6	9025608	15	US-10-156-761-1	Sequence 1, Appli
30	47.5	21.4	746	9	US-09-910-943-530	Sequence 530, App
31	47	21.2	282	12	US-09-864-408A-343	Sequence 343, App
32	47	21.2	831	12	US-10-369-493-35589	Sequence 35589, A
33	47	21.2	955	13	US-10-117-723-524	Sequence 524, App
34	47	21.2	955	15	US-10-037-270-524	Sequence 524, App
35	47	21.2	1128	15	US-10-156-761-2339	Sequence 2339, Ap
36	47	21.2	1140	15	US-10-156-761-1433	Sequence 1433, Ap
37	47	21.2	1191	12	US-10-369-493-34569	Sequence 34569, A
38	47	21.2	1221	13	US-10-214-446-1	Sequence 1, Appli
39	47	21.2	1686	15	US-10-156-761-3852	Sequence 3852, Ap
40	47	21.2	2394	9	US-09-815-242-4016	Sequence 4016, Ap
41	47	21.2	3573	12	US-10-369-493-42233	Sequence 42233, A
42	47	21.2	3933	12	US-10-369-493-32532	Sequence 32532, A
43	47	21.2	4175	15	US-10-307-019-2	Sequence 2, Appli
44	47	21.2	5007	15	US-10-307-019-3	Sequence 3, Appli
45	47	21.2	7893	14	US-10-077-130-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-938-842A-1199
; Sequence 1199, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938.842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1199
; LENGTH: 762
; TYPE: DNA

```

; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1199

Alignment Scores:
Pred. No.: 2,33e-46 Length: 762
Score: 207.00 Matches: 197
Percent Similarity: 100.00% Conservative: 10
Best Local Similarity: 95.17% Mismatches: 0
Query Match: 93.24% Indels: 0
DB: 12 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-938-842A-1199 (1-762)

QY 1 MetSerSerArgTTPAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
DB 1 ATGAGTAGCCGATGGAATCGTACGATCTACGTTGGGAATTTGCCCTGGAGATATTCGCAAG 60
QY 21 CysGluValGluAspLeuPheTyrIleTyrGlyProIleValAspIleAspLeuLysIle 40
DB 61 TGTGAGGTTGAAGATCTCTTCTACAAAGTATGGACCAATTTGGACATTTGAAGATT 120
QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
DB 121 CCACCGAGACCTCTCGTTATGCCCTTGTGAGATCTCTGATCGACGAT 180
QY 61 AlalleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
DB 181 GCAATTTATGACGATGATGATTTGATGCGGTGCTGACCGGATTTACCGCTTCTGCT 240
QY 81 HisGlyGlyArgArg*****SerTyrSerAlaSer 100
DB 241 CATGGTGTGCTAGATTTTCCACATCAGTTGATAGGTACAGCAGCAGCTACAGTGGCAGC 300
QY 101 ArgAlaProSerArgSerArgSerArgValLeuValThrGlyLeuProProSerAla 120
DB 301 CGTGACACCTTCAAGACGCTCTGACTACCGCGTGTGTGACCGGATTTACCGCTTCTGCT 360
QY 121 SerTyrGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
DB 481 TACGCAATTAAGGAACCTTGAATGCGGCTTGTGGATTTAGCAACTATGATGATATGAAG 540
QY 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
DB 541 CGGGTGGGGATATGATCGAGGATGTGAGTCGAAGCCAGATGATTTCTAAAGCTAT 600
QY 201 ArgSerArgSerArgSerArg 207
DB 601 AGAAGCAGGAGTCGGAGCCGT 621

RESULT 2
US-09-938-842A-1199
; Sequence 1199, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24

```

```

; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1199
; LENGTH: 762
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1199

Alignment Scores:
Pred. No.: 2,33e-46 Length: 762
Score: 207.00 Matches: 197
Percent Similarity: 100.00% Conservative: 10
Best Local Similarity: 95.17% Mismatches: 0
Query Match: 93.24% Indels: 0
DB: 12 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-938-842A-1199 (1-762)

QY 1 MetSerSerArgTTPAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
DB 1 ATGAGTAGCCGATGGAATCGTACGATCTACGTTGGGAATTTGCCCTGGAGATATTCGCAAG 60
QY 21 CysGluValGluAspLeuPheTyrIleTyrGlyProIleValAspIleAspLeuLysIle 40
DB 61 TGTGAGGTTGAAGATCTCTTCTACAAAGTATGGACCAATTTGGACATTTGAAGATT 120
QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
DB 121 CCACCGAGACCTCTCGTTATGCCCTTGTGAGATCTCTGATCGACGAT 180
QY 61 AlalleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
DB 181 GCAATTTATGACGATGATGATTTGATGCGGTGCTGACCGGATTTACCGCTTCTGCT 240
QY 81 HisGlyGlyArgArg*****SerTyrSerAlaSer 100
DB 241 CATGGTGTGCTAGATTTTCCACATCAGTTGATAGGTACAGCAGCAGCTACAGTGGCAGC 300
QY 101 ArgAlaProSerArgSerArgSerArgValLeuValThrGlyLeuProProSerAla 120
DB 301 CGTGACACCTTCAAGACGCTCTGACTACCGCGTGTGTGACCGGATTTACCGCTTCTGCT 360
QY 121 SerTyrGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
DB 361 TCGTGCAAGACCTTAAAGGATCACATGCGCAAGCTGGAGATGCTCTGCTTCTTGAAGTT 420
QY 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
DB 421 TTCCTTGACCGTAAAGGCATGCTGGGTTGTGGATTTATAGCAACTATGATGATATGAAG 480
QY 161 TyrAlalleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
DB 481 TACGCAATTAAGGAACCTTGAATGCGGATTTGCAAAATGCTTCTTCTAGTCTTATATA 540
QY 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
DB 541 CGGGTGGGGATATGATCGAGGATGTGAGTCGAAGCCAGATGATTTCTAAAGCTAT 600
QY 201 ArgSerArgSerArgSerArg 207
DB 601 AGAAGCAGGAGTCGGAGCCGT 621

RESULT 3
US-10-006-285-513
; Sequence 513, Application US/10006285
; Publication No. US20030165854A1
; GENERAL INFORMATION:
; APPLICANT: Mary Jane Cunningham
; APPLICANT: Matthew R. Kaser
; TITLE OF INVENTION: MARKER GENES RESPONDING TO TREATMENT WITH TOXINS

```

FILE REFERENCE: PA-0039 US
CURRENT APPLICATION NUMBER: US/10/006,285
CURRENT FILING DATE: 2001-12-05
NUMBER OF SEQ ID NOS: 514
SOFTWARE: PERL Program
SEQ ID NO 513
LENGTH: 3299
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030165854A1 411426.42
US-10-006-285-513

Alignment Scores:
Pred. No.: 1,59e-11 Length: 3299
Score: 95.00 Matches: 108
Percent Similarity: 56.46% Conservative: 10
Best Local Similarity: 51.67% Mismatches: 85
Query Match: 42.79% Indels: 6
DB: 13 Gaps: 2

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-10-006-285-513 (1-3299)

QY 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
DB 203 ATCTAGTGGGTAACTTACCTCCAGACATCCGAACCAAGACATTGAGGACGTCTCTAC 262
QY 29 LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAla 48
DB 263 AATATCGGCGCTATCCGCGACATCGACCTCAAGATCGCGGGGACCGCCCTTCGCC 322
QY 49 PheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyr 68
DB 323 TTCTGTGAGTTCCAGGACCCGCGAGACGCGGAAGACGCGGTGTATGTCGACGCGTAT 382
QY 69 AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArg-----Arg 85
DB 383 GATTACGATGGTACCGTCTCGCGGTGGAGTTTCTCGAAGCGCGCGTGGACAGCGCA 442
QY 86 *****SerTyrSerAlaSerArgAlaProSerArg 105
DB 443 GCGCGGGCGGGGTGGAGGTGGCGGAGTCCCGAGTCCGCTATATGCTGATGTTTACCGAGT 502
QY 106 ArgSerAspTyrArgValLeuValThrGlyLeuProSerAlaSerTrpGlnAspLeu 125
DB 503 CGGTCTGAAACACAGAGTGGTGTCTCTGGACTCCCTCAAGTGGAGTGGCAGGATTGA 562
QY 126 LysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLys 145
DB 563 AAGATCACATGCGTGAAGCAGGTGATGATGATGATGATGATGATGATGATGATGATGAT 616
QY 146 GlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLys 165
DB 617 ---GGCACTGGTCTGGTGGAGTTGTACGGAAGAGATATGACCTATGACGATTCGAAA 673
QY 166 LeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyr 185
DB 674 CTGGATAACACTAAGTTTAGATCTCATGAGGAGAAACTCCCTACATCCGGGTAAAGTT 733
QY 186 GluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArg 205
DB 734 GATGGGCCCAAGATCCAAAGTTATGGAAGATCTCGATCTCGAAGCGGTATGCTAGCAGA 793
QY 206 SerArgGlyProSerCysSerTyrSer 214
DB 794 AGCGGTAGCAGACACACAGGAGT 820

RESULT 4
US-09-880-107-3342
Sequence 3342, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:

APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3342
LENGTH: 1069
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 U30825
US-09-880-107-3342

Alignment Scores:
Pred. No.: 3,23e-11 Length: 1069
Score: 87.50 Matches: 90
Percent Similarity: 48.54% Conservative: 10
Best Local Similarity: 43.69% Mismatches: 101
Query Match: 39.41% Indels: 5
DB: 10 Gaps: 1

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-880-107-3342 (1-1069)

QY 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
DB 98 ATCTAGTGGGTAACTTCCGACCGACGTCGCGAGAGGACTTGGAGGACCTGTTCTAC 157
QY 29 LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAla 48
DB 158 AAGTACGCGCGCATCCGCGACATCGAGTCAAGAACCGCGACCGCCCTCGTCCCTTCGCC 217
QY 49 PheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyr 68
DB 218 TTCGTGCGCTTCGAGGACCCCGGAGATGCGAGAGTCTATTTATGAAGAAATGGTTAT 277
QY 69 AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArg***** 88
DB 278 GATTATGCCAGTGTGGCTTCGTTGTGAG-----TTCCCCAGGACTTAT 322
QY 89 *****SerTyrSerAlaSerArgAlaProSerArgSerArg 108
DB 323 GAGGTGCGGGTGGGTGGCGCCCGTGGGAGGAATGGCGCTCTACAGAAGATCTGAT 382
QY 109 TyrArgValLeuValThrGlyLeuProSerAlaSerTrpGlnAspLeuLysAspHis 126
DB 383 TTCGAGTCTTGTGTTTTCAGGACTTCCCTCGCGACGAGCTGGCAGGACCTGAAGGATCAC 442
QY 129 MetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSer 148
DB 443 ATGCGAGAAGTCTGGGATGCTGTTATGCTGATGTCGAGAAGATGAGTGGGATGGTC 502
QY 149 GlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAla 168
DB 503 GAGTATCTCAGAAAGAAGACATGGAATATCCCTCGGTAACCTGGATGACACCAATTC 562
QY 169 ThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArg 188
DB 563 CGCTCTCATGAGGGTGAACCTTCTCATCCGAGTTTATCTCTGAGAGAAGCACCAGTAT 622
QY 189 SerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGly 208
DB 623 GGTACTACCGGTCTGGGTCTGGGTCAAGGGCGCGTACTCTCCATACCAAGAGCGGGT 682
QY 209 ProSerCysSerTyrSer 214

Db 683 TCCCCACACTACTTGTCT 700

RESULT 5

US-09-925-300-661

; Sequence 661, Application US/09925300

; Patent No. US20020151681A1

; GENERAL INFORMATION:

; APPLICANT: Craig Rosen,

; APPLICANT: Steve Ruben

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PAL01

; CURRENT APPLICATION NUMBER: US/09/925,300

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05988

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 1890

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 661

; LENGTH: 1162

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1155)

; OTHER INFORMATION: n equals a,t,g, or c

US-09-925-300-661

Alignment Scores:

Pred. No.:	4,44e-11	Length:	1162
Score:	87.50	Matches:	90
Percent Similarity:	48.54%	Conservative:	10
Best Local Similarity:	43.69%	Mismatches:	101
Query Match:	39.41%	Indels:	5
DB:	10	Gaps:	1

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-925-300-661 (1-1162)

Qy	9	IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr	28
Db	143	ATCTACGTGGGGAACCTTCCGACCCAGCTGCGCAGAGAGACTTGGAGGACCTGTTCTAC	202
Qy	29	LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAla	48
Db	203	AAGTACGGCCGCATCCGCGAGATCGAGCTCAAGAACCCGCGACGCGCTCGTGCCTTCGCC	262
Qy	49	PheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyr	68
Db	263	TTCGTGCGCTTCGAGGACCCCGAGATCGAGAGGATGCTATTATGGAAGAAATGTTTAT	322
Qy	69	AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArgArg	88
Db	323	GATTATGGCCAGTGTGCGGCTCTGTTGGAG-----TTCCCAGGACTTAT	367
Qy	89	*****SerTyrSerAlaSerArgAlaProSerArgArgSerAsp	108
Db	368	GGAGTCGGGTGGGTGGCCCGTGGTGGGAGAAATGGCCCTCTACAGAGATCTCAT	427
Qy	109	TyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHis	128
Db	428	TTCGAGTTCCTGTTTCAGGACTTCTCCGTCAGCAGCTGGCAGGACCTGAGGATCAC	487
Qy	129	MetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgGlyGlyMetSer	148
Db	488	ATGCAGAAAGCTGGGGATGTCCTGTTATGCTGTCAGAAAGGATGGAGTGGGATGTC	547
Qy	149	GlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAla	168
Db	548	GAGTATCTCAGAAAAGAGACATGGAAATATGCCCTGGCGTAACTGGATGACACCAATTC	607
Qy	169	ThrGluPheArgAsnAlaPheSerSerAlaIleValLeuValLeuValGluSerArg	188

Db 986 ATCGGAGAGCTGGGATCTCTGTTATGCTGATGTCAGAGGATGGAGTGGGGATGCT 927
 QY 149 GlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAla 168
 Db 926 GAGTATCTCAGAAAAGAACATGGAATATGCCCTGCCGTAAACTGGATGACACCAATTC 867
 QY 169 ThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArg 188
 Db 866 CGCTCTCATGAGGTGAACCTTCCTACATCCAGGTTTATCCTGAGAGAGACACGACTAT 807
 QY 189 SerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgGly 208
 Db 806 GGCCTACTCAGCTCTCGGCTCGGTCAGGGCCGCTGACTCTCCATACCAAGCAGGGGT 747
 QY 209 ProSerCysSerTyrSer 214
 Db 746 TCCACACTACTTCTCT 729

RESULT 7
 US-10-171-581-9
 ; Sequence 9, Application US/10171581
 ; Publication No. US20030104426A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dai, Hongyue
 ; APPLICANT: Linsley, Peter
 ; APPLICANT: Mao Mao
 ; TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia
 ; FILE REFERENCE: 9301-157-999
 ; CURRENT APPLICATION NUMBER: US/10/171,581
 ; CURRENT FILING DATE: 2002-06-14
 ; PRIOR APPLICATION NUMBER: 60/298,914
 ; PRIOR FILING DATE: 2001-06-18
 ; NUMBER OF SEQ ID NOS: 366
 ; SEQ ID NO 9
 ; LENGTH: 1717
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: M72709
 ; DATABASE ENTRY DATE: 2001-06-18
 US-10-171-581-9

Alignment Scores:
 Pred. No.: 1,96e-10 Length: 1717
 Score: 87.50 Matches: 89
 Percent Similarity: 49.75% Conservative: 10
 Best Local Similarity: 44.72% Mismatches: 97
 Query Match: 39.41% Indels: 3
 DB: 15 Gaps: 1

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-10-171-581-9 (1-1717)
 QY 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
 Db 148 ATCTAGCTGGTAACTTACCTCCAGACATCCGACCAAGACATTCGAGGACGTGTTCTAC 207
 QY 29 LysTyrGlyProIleValAspIleAspLeuIleProProArgProGlyTyrAla 48
 Db 208 AATACGGCGCTATCCGCGACATCGACCTCAAGATCGCGGGGAGCGCCCTTCGCC 267
 QY 49 PheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyr 68
 Db 268 TTGTTGAGTTTCAGGACCCGCGAGACGCGGTGATGTTGTCGACGCGTAT 327
 QY 69 AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyArg-----Arg 85
 Db 328 GATTACGATGGTACCTCTCGGGTGGAGTTTCTCGAAGCGCGCTGGAACAGGCCGA 387
 QY 86 *****SerTyrSerAlaSerArgAlaProSerArg 105
 Db 388 GCGCGCGGGGTGGAGTGGCGGAGCTCCCGAGTGCCTATGGCCCCCATCCAGG 447
 QY 106 ArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeu 125

Db 448 CGGTCTGAAAACAGAGTGGTGTCTCTGGACTCCCTCCAGTGGAAAGTTTGGCAGGATTTA 507
 QY 126 LysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLys 145
 Db 508 AAGGATCACATGGTGAAGCAGGTGATGTATGTATGCTGATGTTTACCGAGATGGCACT 567
 QY 146 GlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLys 165
 Db 568 GGTGCTCGAGTGTTCGGAAGAAGATATACCTATGCAGTTCGAAAACCTGGATAAC 627
 QY 166 LeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyr 185
 Db 628 ACTAAGTTTAGATCTCATGAGGTAGGTATATACAGTATTCTTTCTTTGACCAGAAATTGG 687
 QY 186 GluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSer 204
 Db 688 ATACAGTGGTCTTAACAGTGGAAATTTTCAGGTAAAGATTTCAGGCAAGTTTGTCAAGT 744

RESULT 8
 US-09-770-791-783
 ; Sequence 783, Application US/09770791
 ; Patent No. US20020062014A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gorlach, Jorn
 ; APPLICANT: An, Yong-Qiang
 ; APPLICANT: Hamilton, Carol M.
 ; APPLICANT: Price, Jennifer L.
 ; APPLICANT: Raines, Tracy M.
 ; APPLICANT: Yu, Yang
 ; APPLICANT: Rameaka, Joshua G.
 ; APPLICANT: Page, Amy
 ; APPLICANT: Matthew, Abraham V.
 ; APPLICANT: Ledford, Brooke L.
 ; APPLICANT: Woessner, Jeffrey P.
 ; APPLICANT: Haas, William David
 ; APPLICANT: Garcia, Carlos A.
 ; APPLICANT: Kricker, Maja
 ; APPLICANT: Slader, Ted
 ; APPLICANT: Davis, Keith R.
 ; APPLICANT: Allen, Keith
 ; APPLICANT: Hoffman, Neil
 ; APPLICANT: Hurban, Patrick
 ; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
 ; FILE REFERENCE: 2029 (PARA-018PRV)
 ; CURRENT APPLICATION NUMBER: US/09/770,791
 ; CURRENT FILING DATE: 2001-01-26
 ; PRIOR APPLICATION NUMBER: 60/178,480
 ; NUMBER OF SEQ ID NOS: 999
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 783
 ; LENGTH: 345
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-770-791-783

Alignment Scores:
 Pred. No.: 3,32e-11 Length: 345
 Score: 81.00 Matches: 71
 Percent Similarity: 77.14% Conservative: 10
 Best Local Similarity: 67.62% Mismatches: 24
 Query Match: 36.49% Indels: 0
 DB: 9 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-770-791-783 (1-345)
 QY 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyValAsnLeuProGlyAspIleArgLys 20
 Db 4 ATGAGCAGTCTGTCAGTAGAACCGGTGACGTCCGAAACCTTCTCGGATATCCGTGAG 63
 QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40

```
Db 64 AGAGAGTCGAAGATTGTTTCAGTAGTATGGACCTGTTTCAAAATTGATTGAAGGTT 123
Qy 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60
Db 124 CCTCAAGGCTCTGTTGTTATGATTCGTTGATTTGATGATCTCGGATGCTGAAGAT 183
Qy 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
Db 184 GCTATTCACTGTTGTTGATGCTATGACTTTGATGGGATCGTTTGAGGTTGGAATTGGCG 243
Qy 81 HisGlyGlyArgArg*****SerTyrSerAlaSer 100
Db 244 CATGTTGGGAGCGTTTCATCAGATGATACTCGGGTAGTTTCAATGGTGGTGGCGGTGT 303
Qy 101 ArgAlaProSerArg 105
Db 304 GGTGGTGGCGCGGT 318
RESULT 9
US-10-060-036-52
; Sequence 52, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Lodes, Michael H.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-036-52
Alignment Scores:
Pred. No.: 1,3e-08 Length: 446
Score: 73.50 Matches: 76
Percent Similarity: 65.15% Conservative: 10
Best Local Similarity: 57.58% Mismatches: 41
Query Match: 33.11% Indels: 5
DB: 15 Gaps: 1
US-10-014-927-19MOD_COPY_1_222 (1-222) x US-10-060-036-52 (1-446)
Qy 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
Db 64 ATCTACCTGGGAGCACTTCGACCGACGTCGCGGAGAGACTTTGAGGACCTGTCTAC 123
Qy 29 LysTyrGlyProIleValAspIleAspLeuIleProProArgProProGlyTyrAla 48
Db 124 AAGTACGCGCGCATCCCGAGATCGAGCTCAAGAACCGGACCGCTCGCTCCCTCGCC 183
Qy 49 PheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyr 68
Db 184 TTCGTGGCTTCGAGGACCCCGAGATGCGAGAGATGCTATTATGAAGAATGGTTAT 243
Qy 69 AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyArgArg***** 88
Db 244 GATTATGGCCAGTGTGGCTTCGTGTGGAG-----ATCCCGAGGACTTAT 288
Qy 89 *****SerTyrSerAlaSerArgAlaProSerArgSerArg 108
Db 289 GGAGGTGGGGTGGGCGCCCGTGGTGGGAGGAATGGCGCTCTCAAGAAGATCTGAT 348
```

```
Qy 109 TyrArgValLeuValThrGlyLeuProSerAlaSerTrpGlnAspLeuLysAspHis 128
Db 349 TTCGAGTCTTGTTCAGGACTTCCTCCGTCCAGCAGCTGGCAGGACCTGAAGATCAC 408
Qy 129 MetArgLysAlaGlyAspValCysPheSerGluVal 140
Db 409 ATCGAGAGAAGCTGGGATGCTGTATTATCTGATGTG 444
RESULT 10
US-09-910-943-651
; Sequence 651, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Briavanlou, Ali
; APPLICANT: Altmann, Curtis
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7529/1G148US1
; CURRENT APPLICATION NUMBER: US/09/910,943
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 742
; SOFTWARE: Patent version 3.1
; SEQ ID NO 651
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Xenopus laevis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(759)
; OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-651
Alignment Scores:
Pred. No.: 1.98e-06 Length: 759
Score: 69.00 Matches: 81
Percent Similarity: 42.92% Conservative: 10
Best Local Similarity: 38.21% Mismatches: 117
Query Match: 31.08% Indels: 4
DB: 9 Gaps: 2
US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-910-943-651 (1-759)
Qy 2 SerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCys 21
Db 109 TCTGGATCCGGCGATGGACGGATATATGTCGGGAATTCGCCGTCTGATATTCGGGAGAAG 168
Qy 22 GluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLys---Ile 40
Db 169 GAAGTACAGAGATCTCTTTGATCGCTATGGTAGGATCCGGACCGTAGAGTTGAAGAACGG 228
Qy 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60
Db 229 GCGGCGAGTAGTGGCCCATTCGCATTTCATCAGCTATCAGGACCCCGCTGATCGAGAGGAC 288
Qy 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGlu----- 78
Db 289 GCAGTGTTCGGAGGAATGGCTATGACTTTGGCTCGTGTGCTGCTACGTGTGATTTCCG 348
Qy 79 ---IleAlaHisGlyArgArg*****SerTyr 97
Db 349 GTTCTTCGGGATCTCGCGCGGAGGTGGTGGCGGGGATATGAGGCTCCCG 408
Qy 98 SerAlaSerArgAlaProSerArgSerAspTyrArgValLeuValThrGlyLeuPro 117
Db 409 GGAAGAATGTCGCCCATCTCGCGCTCTGAATACAGAGTCATTGTCTCAGGCTCTTCCA 468
Qy 118 ProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPhe 137
Db 469 CCCTCAGAGAGCTGGCAGGATCTGAGGATCATATGCGGAAGCTGGTGTGATGTCTGTAT 528
Qy 138 SerGluValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAsp 157
Db 529 GCTTGATGTACACAAGATGGAAATGGGATAGTCGAATTCATTCACAAAGAGATATGGAA 588
```



```

RESULT 14
US-10-388-934-197
; Sequence 197, Application US/10388934
; Publication No. US20040005547A1
; GENERAL INFORMATION:
; APPLICANT: Boess, Franziska
; APPLICANT: Suter-Dick, Laura
; APPLICANT: Wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 197
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US20040005547Alway rat)
US-10-388-934-197

Alignment Scores:
P-Ided. No.:          Length:      1579
Score:                Matches:       47
Percent Similarity:   26.03%
Best local Similarity: 21.46%
Query Match:         25.68%
DB:                  Indels:        0
                       Gaps:         0

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-10-388-934-197 (1-1579)
Qy    Qy           3 SerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArglysCysGlu 22
Db    Db           243 AGTGCGCTGCAGTGTTTCATTGGGACATAAAATCCCGCAGCGAGGAGGAAGAAGATGTGGAA 302
Qy    Qy           23 ValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProPro 42
Db    Db           303 AGATTCITCAAGGTTCAGCGCGCATCAGAGATATTGATTGMAAAGAGGTTTTGGTTTT 362
Qy    Qy           43 ArgProProGlyTyraIlaPheValGluPheGluaspProArgAspAlahaspaspaIalle 62
Db    Db           363 GTGGAAITTAGGCCACCACGGGATGCAGATGACGCTGTTTATGAACCTTGATGGGAAGGAA 422
Qy    Qy           63 TyrGlyArgaspGlyTyraSpPheaspGlyCysArgLeuArgValGluIlealaHisGly 82
Db    Db           423 CTTTGCAGTAAGGGTTACTATTAAACAATGCTCGGGCTCGGTCTCGAGGTGGAAAGAGGT 482
Qy    Qy           83 GlyArgarg*****SerTyrSerAlaserargalla 102
Db    Db           483 AAGAGCACGACTACCGACCGTTTTAGCAGTCGCGAGACCTCGAAAATGATAGACGAATGCT 542
Qy    Qy           103 ProSerArgSerpTyrArgValLeuValThrGlyLeuProProSerAlaserTp 122
Db    Db           543 CCACCTGAAGACAGAAAAATCGACITATAGTTGCAAATTTATCCTCAAGAGTCAGCTGG 602
Qy    Qy           123 GlNaspLeuylsaspHismetArgLysAlaGlyAspValCyepheserGluvalPhePro 142
Db    Db           603 CAGGATCTCAAAAGATTTCATGACAGAACGCTGGGAAAGTAACCTTTTCGGATGCAATCGA 662
Qy    Qy           143 AspargylsglyMetserGlylvaValaspTyrSerAsnTyrYraspMetylstyrala 162
Db    Db           663 CCTAAACTAAATGAAGGGGTAGTTGAGTTGGCTCTCTTTATGGTGACCTTAAGAATGCTATT 722
Qy    Qy           163 IleArgLysLeuaspaIatThrGluPheArgasnAlapheSeSerAlatyrlileArgval 182
Db    Db           723 GAGAAACTTCTCGAAGGAAATTAATGGGAGHAATAATCAATTAATTGAAGCAGCAAA 782
Qy    Qy           183 ArgGluTyrGluserArgSerValSerArgSerProAspAspSerLysserTyrArgSer 202

```

Db 783 AGACACAGGTCAGAGCAGGTCACGATCTCGGACCGAGGTTCTCTAGGTCGCGTAGC 842
 Qy 203 ArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSer 221
 Db 843 CGATCCCGTTCAGCGAGGACGAAGTCTTACAGCCGATCAAGGAGCAGGAGCGCGAGC 899

RESULT 15

US-10-014-927-18
 ; Sequence 18, Application US/10014927
 ; Publication No. US20020115180A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Barta, Andrea
 ; APPLICANT: Lepato, Sergey
 ; APPLICANT: Kalyna, Maria
 ; APPLICANT: Dornier, Silke
 ; TITLE OF INVENTION: Splice Factor
 ; FILE REFERENCE: SONN:013US
 ; CURRENT APPLICATION NUMBER: US/10/014, 927
 ; CURRENT FILING DATE: 2001-10-23
 ; PRIOR APPLICATION NUMBER: PCT/AT00/00100
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: A 727/99
 ; PRIOR FILING DATE: 1999-04-23
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 18
 ; LENGTH: 4044
 ; TYPE: DNA
 ; ORGANISM: Unknown
 ; FEATURE:
 ; OTHER INFORMATION: Description of the unknown organism: genome
 ; OTHER INFORMATION: atSRp30
 US-10-014-927-18

Alignment Scores:
 Pred. No.: 6.67 Length: 4044
 Score: 56.00 Matches: 46
 Percent Similarity: 26.92% Conservative: 10
 Best Local Similarity: 22.12% Mismatches: 152
 Query Match: 25.23% Indels: 0
 DB: 14 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-10-014-927-18 (1-4044)

Qy 11 ValGlyAsnLeuProGlyAspLeuArgLysCysGluValGluAspLeuPheTyrLysTyr 30
 Db 1659 GTGATTAGGATCACATCGCGAAGCTGGAGATGTCGTCTCTGAAGTTTCCCTGACC 1718
 Qy 31 GlyProileValAspLeuLysLleProProArgProGlyTyrAlaPheVal 50
 Db 1719 GTAAAGGTGAGTTGACATTCGATAGTTTCGATAAGCTTTTTCGATGATGTTAGTAAT 1778
 Qy 51 GluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyrAspPhe 70
 Db 1779 TAGTCTTTGTGAAGGAGATAGTGTAAAGCATCTGAACCTGTAACCTCACATTCAGTAT 1838
 Qy 71 AspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArg***** 90
 Db 1839 TTCCTTTGAGGCATGTCGGGTTGCGATTATAGCAACTATGATGATGAAGTACGCA 1898
 Qy 91 *****SerTyrSerAlaSerAlaProSerArgArgSerAspTyrArg 110
 Db 1899 GTAAGTTTATATCTTTGCAACGCAATGTTCTCTGGAGCCAGTGTGATCTTTGATGTC 1958
 Qy 111 ValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArg 130
 Db 1959 TTCATAGTATACCGACCTGAATTTATCTTCTCTGGAGCCAGTGTGATCTTTGATGTC 2018
 Qy 131 LysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyVal 150
 Db 2019 CCTTAAATTTTGTGATGTGACAGATAAGGAAACTTGTATGCCACTGAATTCGAAATGCTT 2078
 Qy 151 ValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGlu 170

Db 2079 TCCTAGTGTCTATATACGGGTATGTTGTATTGCTTTCTTTGATTTTGTAAAGCATAGT 2138
 Qy 171 PheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArgSerVal 190
 Db 2139 GGATATGGAGTCATCTCTGAATTTACTGTTTCAGGTGAGGGAATATGATGTCGAGGAGTGTG 2198
 Qy 191 SerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGlyProSer 210
 Db 2199 AGTCGAAGCCAGATGATTTCTAAAGCTATAGAAGCAGGAGTCGGAGCCGTGTGTCACAGC 2258
 Qy 211 CysSerTyrSerSerLysSerArg 218
 Db 2259 TGTAGCTATAGTAGCAAGACAGCAGG 2282

Search completed: February 4, 2004, 21:32:39
 Job time : 357 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2004, 21:15:14 ; Search time 2041 Seconds

(without alignments)
2643.603 Million cell updates/sec

Title: US-10-014-927-19mod_copy_1_222

Perfect score: 222

Sequence: 1 MSSRWNTIYVGNLPGDIRK.....RSRSRGSCSYSKSRVSP 222

Scoring table: UNITARY2

Xgapop 10.0 , Xgapext 0.5

Fgapop 10.0 , Fgapext 0.5

Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-C=/cgn2_1/USPTO.spool/BAUM927/runat_04022004_131334_2716/app_query.fasta_1.391
-DB=5ST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=unitary2 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THRESHOLD=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=BAUM927 -CGEN 1 1 2810 @runat_04022004_131334_2716 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estio:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	160	72.1	604	9	AV826310	AV826310 AV826310
2	156	70.3	779	13	BQ802976	BQ802976 WHE2832 C
3	154	69.4	647	12	BI931752	BI931752 EST551541
4	153	68.9	877	10	BG444501	BG444501 GA_Ea002
5	152	68.5	627	10	BG135885	BG135885 EST468777
6	151	68.0	638	13	BQ986483	BQ986483 HFI1P24r
7	151	68.0	686	10	BG440250	BG440250 GA_Ea000
8	150	67.6	834	9	AW448238	AW448238 BRY_1489
9	150	67.6	834	13	BQ605903	BQ605903 BRY_1489
10	150	67.6	890	10	BG414461	BG414461 HVSMEK000
11	148	66.7	643	12	BM111827	BM111827 EST559363
12	147	66.2	852	14	CD439565	CD439565 EL01N0526
13	146.5	66.0	1255	11	AY108591	AY108591 Zea mays
14	146	65.8	614	14	CA008486	CA008486 HUI1B23r
15	146	65.8	625	12	BJ269490	BJ269490 BJ269490
16	146	65.8	652	12	BQ211859	BQ211859 BJ211859
17	146	65.8	687	13	BQ841884	BQ841884 WHE2986 D
18	145	65.3	631	13	BQ240790	BQ240790 Ta05012H
19	145	65.3	658	12	BJ14572	BJ14572 BJ14572
20	145	65.3	667	10	BG451868	BG451868 NF101A10D
21	144.5	65.1	840	14	CB650587	CB650587 OSUNE15B
22	143	64.4	560	10	BG134853	BG134853 EST467745
23	143	64.4	629	9	AI482838	AI482838 EST242161
24	143	64.4	666	12	BM878853	BM878853 P5-G02 Sw
25	143	64.4	736	14	CA802008	CA802008 sau28e1.0
26	143	64.4	741	14	CB635325	CB635325 OS11EB15J
27	143	64.4	778	10	BF274242	BF274242 GA_Eb002
28	142.5	64.2	810	14	CB662124	CB662124 OSUNE05N
29	142	64.0	604	12	BJ209636	BJ209636 BJ209636
30	142	64.0	640	14	CA018302	CA018302 HV08D15r
31	142	64.0	813	14	CB634299	CB634299 OS11EB13L
32	141.5	63.7	677	12	BM063881	BM063881 KS01060F0
33	141	63.5	604	12	BI179535	BI179535 EST520480
34	141	63.5	657	12	BJ302592	BJ302592 BJ302592
35	141	63.5	694	10	BF622642	BF622642 HVSME000
36	141	63.5	719	14	CA783335	CA783335 sat22g12
37	140	63.1	612	10	BE404280	BE404280 WHE1204 D
38	140	63.1	638	13	BQ238916	BQ238916 TaE05039F
39	140	63.1	675	10	BG648263	BG648263 EST509882
40	139	62.6	677	9	AL508965	AL508965 AL508965
41	138.5	62.4	641	12	BM099586	BM099586 EBes01 SQ
42	137.5	61.9	624	10	BG368538	BG368538 HVSME101
43	137.5	61.9	667	10	BE420047	BE420047 WWS02.86R
44	137	61.7	599	14	CA012262	CA012262 HT04019r
45	137	61.7	654	12	BJ446900	BJ446900 BJ446900

ALIGNMENTS

RESULT 1
AV826310
LOCUS AV826310 RAPL8 Arabidopsis thaliana cdna clone RAPL08-12-105 5',
DEFINITION AV826310 mRNA linear EST 01-APR-2002
mRNA sequence.
ACCESSION AV826310
VERSION AV826310.1 GI:19868370
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

US-10-014-927-19MOD_COPY_1_222 (1-222) x BQ802976 (1-779)

```
Qy 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
Db 14 ATGAGTAGGCGCTGGAGCCGACCAATTACGTTGGGAACCTCCAGGGGATATCAGGAG 73
Qy 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLys 40
Db 74 CGGAGGTGGAGGATCTCTTCAAGATATGGACGATAGTTGAAATTGACTTGAAGGTC 133
Qy 41 ProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
Db 134 CCCCCAAGGCTCTGGCTTGTCTTTGTTGAGTTTGAAGATCCCGCTGATGCTGAGAT 193
Qy 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluLeuAla 80
Db 194 GCGATTATGCGCCAGATGCTACAACTTTGATGGGAATCGGCTTAGGCTTGAACCTTGA 253
Qy 81 HisGlyGlyArgArg*****SerTyrSerAlaSer 100
Db 254 CATGGGGGAGGCGCAACTCTTCATCCCTCCGACAGCTATGTTGGTGGGGAGCGCGCT 313
Qy 101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProSerAla 120
Db 314 GTGGTGTCTCTAGGATACGGAGTATCGTGTCTGTACTGAGTACTCTTCTCTGCA 373
Qy 121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
Db 374 TCATGGCAAGATCTAAGGACCATATGAGGAAGCGCTGGTGTGTTGTTCTCTGAGGTG 433
Qy 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
Db 434 TACCGTGAGGGTGGTGTACTATTGGAATTTGTTGATTATACAACTATGATGATGAAG 493
Qy 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
Db 494 TATGCTATAGGAAGCTTGATGATGATCTGAATTTTAAATGCTCTCTCGAGCGCCTATA 553
Qy 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
Db 554 AGGGTGAGGAGATGCTGGAAGAACGAGCGGCTCTCTATTCAGCAGCGGTAGCAGAAC 613
Qy 201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220
Db 614 CGAAGTGGAGCTACAGCAGAGTCCAGTCCAAAGAAACCAACATCAGCGGCTCAGCG 673
Qy 221 Ser 221
Db 674 TCA 676
```

RESULT 3
BI931752
LOCUS
DEFINITION
EST551641 tomato flower, 8 mm to preanthesis buds Lycopersicon
esculentum cDNA clone CTC21A15 5' end, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

647 bp mRNA linear EST 18-OCT-2001
EST551641 tomato flower, 8 mm to preanthesis buds Lycopersicon
esculentum cDNA clone CTC21A15 5' end, mRNA sequence.
BI931752
BI931752.1 GI:16246224
EST.
Lycopersicon esculentum (tomato)
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 647)
van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
Uttterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
Generation of ESTs from tomato flower tissue, buds 8 mm -
preanthesis
Unpublished
Contact: CUGI
Clemson University Genomics Institute
Clemson University

100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
Location/Qualifiers
source
1..647
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA96"
/db_xref="taxon:4081"
/clone="CTC21A15"
/tissue_type="flower"
/dev_stage="buds 8mm to preanthesis"
/clone_lib="tomato flower, 8 mm to preanthesis buds"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA96).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."
BASE COUNT 166 a 122 c 172 g 187 t
ORIGIN

Alignment Scores:
Pred. No.: 1.72e-27 Length: 647
Score: 154.00 Matches: 144
Percent Similarity: 74.40% Conservative: 10
Best Local Similarity: 69.57% Mismatches: 53
Query Match: 69.37% Indels: 0
DB: 12 Gaps: 0
US-10-014-927-19MOD_COPY_1_222 (1-222) x BI931752 (1-647)

```
Qy 4 ArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluVal 23
Db 27 CGTCTAAGTCGAGTATCTACGTTGGAAATCTCTCGTGATATTCGGGAGAGAGAATA 86
Qy 24 GluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProArg 43
Db 87 GAAATTTGTTTACAGATATGGTCCCATTTGGAATTTGATTGAAATTCACCTAGA 146
Qy 44 ProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyr 63
Db 147 CCACCTGGTTATGCTTCTGTTAGAGTTTGAAGATCCTCGTGATGCTGATGATCGT 206
Qy 64 GlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGly 83
Db 207 GGGGTGATGGCTATGACTTTGACGGGCATCGCTGGAGTTGAACCTTGACATGGTGG 266
Qy 84 ArgArg*****SerTyrSerAlaSerArgAlaPro 103
Db 267 CGAGGATCATCATATGATCGCCACAGTAGTTACAGTAGTCGAGTCGAGTGGACTT 326
Qy 104 SerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGln 123
Db 327 TCTAGGCGCTCTGACTATCGGCTACTGGTCTCTGAGTACGCTCTTCTGCTCTATG 386
Qy 124 AspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAsp 143
Db 387 GACITGAAGGATCATATGCGACGAGCTGGAGATGCTGCTTCTCTCAAGTTTCCGAG 446
Qy 144 ArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIle 163
Db 447 COTGACGGTATGAGAGGGGATTTGGAGCTTATACCACTATGATGATGATGATGATG 506
Qy 164 ArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArg 183
Db 507 AAGAAACTTGTAGTCTCTCTTTTCCGAATCAATTTCTCGAGCATATATTAGGTTG 566
Qy 184 GluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArg 203
Db 567 AAGTATGATAGAGCATAGTATTCAGGAGGTCAGTCCATCACTTCTAGAGCAGA 626
```


QY 204 SerArgSerArgGlyProSer 210 877 bp mRNA linear EST 15-MAR-2001
Db 627 AGGTACTCAAGAGTAGGAGT 647
RESULT 4
Bg444501
LOCUS
DEFINITION GA_Ea0024123f Gossypium arboreum 7-10 dpa fiber library Gossypium
arborescens cDNA clone GA_Ea0024123f, mRNA sequence.
ACCESSION
Bg444501
VERSION
Bg444501.1 GI:13354153
KEYWORDS
EST.
SOURCE
Gossypium arboreum
ORGANISM
Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE
1 (bases 1 to 877)
Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry
D., Wood, T.C., Leslie, A. and Wilkins, T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished
JOURNAL
COMMENT
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCATATAGGG
High quality sequence stop: 706.
Location/Qualifiers
1..877
/organism="Gossypium arboreum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0024123f"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 218 a 194 c 228 g 236 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1,11e-26 Length: 877
Score: 153.00 Matches: 143
Percent Similarity: 71.16% Conservative: 10
Best Local Similarity: 66.51% Mismatches: 62
Query Match: 68.92% Indels: 0
DB: 10 Gaps: 0
US-10-014-927-19MOD_COPY_1_222 (1-222) x BG444501 (1-877)
QY 1 MetSerSerArgTTPAsnArgThrIleTyrValGlyAenLeuProGlyAspIleArgLys 20
Db 90 ATGAGTACCGCTCCAGCAGGACTTTACGTTGGAATCTTCCGGTGATGTCGGCAG 149
QY 21 CysGluValGluAspLeuPheTyrIleTyrGlyProIleValAspIleAspLeuLysIle 40
Db 150 AGGGAAGTGAAGATTTCCTTTTATAAGATGATGTCCTCCATAGTCAATAGTCTGAGATT 209
QY 41 ProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60
Db 210 CCACCAAGCCTCCAGGTATGCAATTTTCAGTTTGAAGATGTCGATGTCGAGAT 269
QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80

Db 270 GCCATTGCTGCTGCTGATGATATGACTTTGGAGGGCATCGTTTACGGGTTGAACCTTGT 329
QY 81 HisGlyGlyArgArg*****SerTyrSerAlaSer 100
330 CATGTGCTGCTGCTGCTGCTTCCATAGATCGTCACAGCAGTTATAGTAGTGGCGGTGA 389
QY 101 ArgAlaProSerArgSerArgSerArgValThrGlyLeuProSerAla 120
390 CGTGGACCATCAGCGCTTCAATATCGGTGCTAGTACTGGATGCCATCTTCTGCT 449
QY 121 SerTTPGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
450 TCATGGCAGGACCTCAGGATCAGATGCGTCCAGCAGGAGATGTTTGTCTCAAGTT 509
QY 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspMetLys 160
510 TTCCTGATGCTAGTGGGACTACAGGATGTTGGACTACACCACTATGATGATGAAG 569
QY 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
570 TATGCTATTAGAAACATCGATGACTGCTGAGTTTCGTAATGCAATTTCTCGGCGCATGTT 629
QY 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
630 CGGGTTAGGAATATGATTCTAGCGGGGATTCCTCTAGAAGCCCTAGTGTGGCGCACT 689
QY 201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSer 215
690 CTCCTAAGAGCCGAGCCGAGTAGAGCCGAGCAAGCCGCTGCTG 734
RESULT 5
Bg135885
LOCUS
DEFINITION .CT02232 5' sequence, mRNA sequence.
ACCESSION
Bg135885
VERSION
Bg135885.1 GI:12636073
KEYWORDS
EST.
SOURCE
Lycopersicon esculentum (tomato)
ORGANISM
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
1 (bases 1 to 627)
van der Hoeven, R., Sun, H., Cho, J., Utterback, T., Hansen, C., Ronning
C. and Tanksley, S.
Generation of ESTs from tomato crown gall tissue
Unpublished
COMMENT
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
Location/Qualifiers
1..627
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CT02232"
/tissue_type="crown gall"
/dev_stage="crown galls from full-grown plants (8 wks old
/lab_host="SOLR"
/clone_lib="tomato crown gall"
XhoI; Four wk old greenhouse plants were stab inoculated
on stem with Agrobacterium tumefaciens C58 (Dr. T.J. Burr,
Cornell U.). Galls were allowed to develop for another 4
wks, when gall tissue was frozen in liquid nitrogen."
BASE COUNT 159 a 123 c 164 g 181 t
ORIGIN

Alignment Scores:

Pred. No.: 5.24e-27 Length: 627
Score: 152.00 Matches: 142
Percent Similarity: 75.62% Conservativity: 10
Best Local Similarity: 70.65% Mismatches: 49
Query Match: 68.47% Indels: 0
DB: 10 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x BG135885 (1-627)

```
Qy 4 ArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspLeuArgLysCysGluVal 23
Db 20 CGTCTAAGTCGCGATATCTAGCTTGGAAATCTTCTCGTGATATTCGGGAGAGTA 79
Qy 24 GluAspLeuPheTyrLysTyrGlyProIleValAspLeuLysIleProProArg 43
Db 80 GAAGATTGTTTACAAAGTATGTCCTCCATTTGGAAATGATTGAAGTTCACCTAGA 139
Qy 44 ProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyr 63
Db 140 CCACCTGGTATGCTTCTAGAGTTGAAGATCTCTGATGCTGATGATGCTATCGT 199
Qy 64 GlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly 83
Db 200 GGGGCTGATGCTATGACTTTGACGGGCATCGCTTGGAGTTGAAGTTCACATGTTGG 259
Qy 84 ArgArg*****SerTyrSerAlaSerArgAlaPro 103
Db 260 CGAGGATCATCATATGATCGGCACAGTAGTACAGTAGTCGAGTCGAGTGGACTT 319
Qy 104 SerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGln 123
Db 320 TCTAGGCGCTGACTATCGCTACTGCTCTGAGACTCGCTCTCTGCTCATGCGCAA 379
Qy 124 AspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAsp 143
Db 380 GACTTGAGGATCATATGCGACGAGCTGGAGATGCTGCTCTCTCAAGTTTCCAGAT 439
Qy 144 ArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspMetLysTyrAlaIle 163
Db 440 CGTGACGGTATGACAGGAGTGTGGACTATACCACTATCATGATGATGATGACGGATA 499
Qy 164 ArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArg 183
Db 500 AAGAACTTGATGACTCTCTGTTTCCGAATCAATCTCTCGACATATATTAGGTGGAC 559
Qy 184 GluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArg 203
Db 560 AAGTATGATAAGAGGATGATGCTATTCAGGAGTCCAAAGTCCATACATTCATAGACGA 619
Qy 204 Ser 204
Db 620 AGT 622
```

RESULT 6

BU986483

LOCUS

DEFINITION HF11P24r HF Hordeum vulgare subsp. vulgare cDNA clone HF11P24

5-PRIME, mRNA sequence.

ACCESSION

BU986483

VERSION

EST.

KEYWORDS

Hordeum vulgare subsp. vulgare

SOURCE

Hordeum vulgare subsp. vulgare

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae

REFERENCE

1 (bases 1 to 638)

AUTHORS

Radchuk, V., Zhang, H., Weesche, W., Potokina, E. and Wobus, U.

TITLE

Barley ESTs from developing seeds

JOURNAL

Unpublished

COMMENT

Contact: Stein Nils

Molecular Markers Group, Department Genbank

Institute of Plant Genetics and Crop Plant Research (IPK)

Cortensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 638 Std Error: 0.00
Plate: 11 row: p column: 24
Seq primer: M13rev.

Location/Qualifiers

1..638

/organism="Hordeum vulgare subsp. vulgare"

/mol_type="mRNA"

/cultivar="barke"

/db_xref="GABI:246423"

/db_xref="taxon:112509"

/clone="HF11P24"

/tissue_type="caryopsis"

/dev_stage="developing caryopsis, 16-25 DAF (days after flowering)"

/lab_host="XL10-Gold"

/clone_lib="HF"

/note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of cDNA); Site 2: XhoI (3'-end of cDNA); developing caryopsis, 16-25 DAF (days after flowering) Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also due to the cloning system used Blue/white selection for recombinants is not 100% reliable. Average insert size is 940 bp"

BASE COUNT 153 a 127 c 195 g 163 t

ORIGIN

Alignment Scores:

Pred. No.: 1.05e-26 Length: 638
Score: 151.00 Matches: 141
Percent Similarity: 72.95% Conservativity: 10
Best Local Similarity: 68.12% Mismatches: 56
Query Match: 68.02% Indels: 0
DB: 13 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x BU986483 (1-638)

```
Qy 4 ArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspLeuArgLysCysGluVal 23
Db 13 CGCTGAGCGGAGCCATTTACGTTGGGAACCTCCAGGGGATATCAGGAGCGGAGGTG 72
Qy 24 GluAspLeuPheTyrLysTyrGlyProIleValAspLeuLysIleProProArg 43
Db 73 GAGGATCTGTTCTACAGTATGCGGATAGTGGAAATTTGACTTGAAGTCCCCCAAG 132
Qy 44 ProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyr 63
Db 133 CCACCTGGCTTTGCTTTTGTGAGTTTGAAGATCCCGTGAAGCGGAGGATCGATTGAG 192
Qy 64 GlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly 83
Db 193 GGCGAGATGCTACACTTTGATGGGATCGCTGAGGGTTGAAGTTCACATGTTGGG 252
Qy 84 ArgArg*****SerTyrSerAlaSerArgAlaPro 103
Db 253 AGGGCTAACTCTTCATCCCTTCTTAACAGCCATGCTGTGTGGAGGAGCGCGTGTGTGTC 312
Qy 104 SerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGln 123
Db 313 TCTAGGCATACAGAGTATCGTGTCTCTGCTTACTGACTTACCTTTCTCTGATCATGCGAA 372
Qy 124 AspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAsp 143
Db 373 GATCTAAAGGATCATATAGAAAGCGTGTGATGTTTGTCTCTGAGGTATCGTGAG 432
Qy 144 ArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspMetLysTyrAlaIle 163
```

Db	433	GGTGGTGTACTATTGGAAATTTGTTGATTATACAAACTATGATGATATGAAGTATGCTATA	492
QY	164	ArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArg	183
Db	493	AGGAAGCTTGATGATACGATGAAATTTAAAATGCTTCTCTCGAGCGCTATAGAAGGTGAAG	552
QY	184	GluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArg	203
Db	553	GAGTATGCTGCCAAAGAGCGCGTCTCTATTCTCGCAGCGGTAGCAGAACCGGAAGTGGC	612
QY	204	SerArgSerArgGlyProSer	210
Db	613	AGCTACAGCAGAGTCCGAGC	633
RESULT 7			
LOCUS	BG440250	686 bp mRNA linear EST 15-MAR-2001	
DEFINITION	GA_Ea0006K16f Gossypium arboreum 7-10 dpa fiber library Gossypium		
ACCESSION	BG440250		
VERSION	BG440250.1	GI:13349901	
KEYWORDS	EST.		
SOURCE	Gossypium arboreum		
ORGANISM	Gossypium arboreum		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids		
AUTHORS	; euroids II; Malvales; Malvaceae; Malvoideae; Gossypium.		
TITLE	Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry, D., Wood, T.C., Leslie, A. and Wilkins, T.A. An integrated analysis of the genetics, development, and evolution of the cotton fiber		
JOURNAL	Unpublished		
COMMENT	Contact: Wing RA Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 7288 Fax: 864 656 4293 Email: rwing@clemson.edu Seq primer: TAATACGACTCACTATAGG High quality sequence stop: 682. Location/Qualifiers		
FEATURES	source		
	1..686	/organism="Gossypium arboreum"	
		/mol_type="mRNA"	
		/strain="AKA"	
		/cultivar="8400"	
		/db_xref="taxon:29729"	
		/clone="GA_Ea006K16f"	
		/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"	
		/lab_host="E. coli"	
		/clone_lib="Gossypium arboreum 7-10 dpa fiber library"	
		/note="Vector: p3K-CMV; Site_1: EcoRI; Site_2: XhoI"	
BASE COUNT	167 a	136 c	184 g
ORIGIN	199 t		
Alignment Scores:			
Pred. No.:	1.41e-25	Length:	686
Score:	151.00	Matches:	141
Percent Similarity:	74.75%	Conservative:	10
Best Local Similarity:	69.80%	Mismatches:	51
Query Match:	68.02%	Indels:	0
DB:	10	Gaps:	0
US-10-014-927-19MOD_COPY_1_222 (1-222) x BG440250 (1-686)			
QY	1	MetSerSerArgTTPAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys	20
Db	79	ATGAGTACCGGTCCAGCAGGACTCTTACGTTGGAATCTCCCGTGATGTTCGGAG	138
QY	21	CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle	40
Db	139	AGGGAAGTGGGAAGATTTGTTTATAGTATGGTCCCATAGCTCAATTTGAGATT	198
QY	41	ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp	60
Db	199	CCACCAAGGCGCTCCAGGTTATGCAITTTGTGAGTTTGAAGAAGCTCGAGATGCTCAAGAT	258
QY	61	AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla	80
Db	259	GCAATTCGTGGTGGTGGATATGACTTTGAGGGCATCGTTTACGGGTTGAATTCCT	318
QY	81	HisGlyGlyArgArg*****SerTyrSerAlaSer	100
Db	319	CATGGTGGTGGGCGGCTCTCCATAGATCGTCACAGCATATAGTAGTGGGCGTGA	378
QY	101	ArgAlaProSerArgSerAspTyrArgValLeuValThrGlyLeuProSerAla	120
Db	379	CGTGGACCATCCAGGGTCTGAATATCGGTGCTAGTTACTGGATTGCCATCTCTGCT	438
QY	121	SerTTPGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal	140
Db	439	TCATGGCAGGACCTCAAGGATCACATGGGTGCAGCAGGAGATGTTTGTGTTTCTCAAGTT	498
QY	141	PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys	160
Db	499	TTCCGTGATGTAGTGGGACTACAGGATTGTGGACTACACCACTATGATGATGAAG	558
QY	161	TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle	180
Db	559	TATGTTATTAGAACTCGATGACTCTGAGTTTCGTAATGATGATTTCTCGGGCATATGTT	618
QY	181	ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr	200
Db	619	CGGGTTAAGGATATGATTCTAGCGGGGATTCCTCTAGAGCCCTAGTCTGGCGGATCT	678
QY	201	ArgSer	202
Db	679	CTCTCA	684
RESULT 8			
LOCUS	AW448238	834 bp mRNA linear EST 03-JAN-2001	
DEFINITION	BRY_1489 BRY Triticum aestivum cDNA clone P52-1A, mRNA sequence.		
ACCESSION	AW448238		
VERSION	AW448238.1	GI:12018666	
KEYWORDS	EST.		
SOURCE	Triticum aestivum (bread wheat)		
ORGANISM	Triticum aestivum		
REFERENCE	1 (bases 1 to 834)		
AUTHORS	Clarke, B.C., Hobbs, M. and Appels, R.		
TITLE	Genes active in developing wheat endosperm		
JOURNAL	Unpublished		
COMMENT	Contact: Bryan Clarke Division of Plant Industry C.S.I.R.O. GPO Box 1600, Canberra, ACT, Australia Tel: 61 2 6246 5054 Fax: 61 2 6246 5000 Email: bryan@pi.csiro.au. Location/Qualifiers		
FEATURES	source		
	1..834	/organism="Triticum aestivum"	
		/mol_type="mRNA"	
		/cultivar="Wyuna"	
		/db_xref="taxon:4565"	
		/clone="P52-1A"	
		/cell_type="endosperm"	
		/clone_lib="BRY"	
BASE COUNT	198 a	199 c	234 g
ORIGIN	200 t		
	3 others		

Alignment Scores: 5.84e-26 Length: 834
Pred. No.: 150.00 Matches: 140
Score: 71.43% Conservative: 10
Percent Similarity: 66.67% Mismatches: 60
Best Local Similarity: 67.57% Indels: 0
Query Match: 67.57% Gaps: 0
DB: 13

US-10-014-927-19MOD_COPY_1_222 (1-222) x AM448238 (1-834)

QY 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
Db 123 ATGAGTAGGCGCTGGAGCGGACCATTTAGTTGGGAACCTCCAGGGGATATCAGGAG 182

QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40
Db 183 CGGAGGTGGAGGATCTCTTCAAGATATGAGCGGATAGTTGAAATTCACTTGAAGTC 242

QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspIleAspAsp 60
Db 243 CCCCCAAGGCTCTCGCTTGTGTTGTTGAGTTTGAAGATCCCGCTGATCTGAAGAT 302

QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuValGluIleAla 80
Db 303 GCATTCATGGCCGAGATGGCTACACTGATGGGAATCGGCTTANGTTGAAGTGA 362

QY 81 HisGlyGlyArgArg*****SerTyrSerAlaSer 100
Db 363 CATGGCGGAGGCGCAACTCTTCATCCCTCCGAACAGCTATGTTGGTGGGAGCGCGT 422

QY 101 ArgAlaProSerArgSerArgSerArgValLeuValThrGlyLeuProSerAla 120
Db 423 GGTGGTGTCTTAGGCATACGGAGTATCGTGTCTTCTGTTACGACCTACCTTCTCTGCA 482

QY 121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
Db 483 TCATGGCAAGATCTGAAGGACCATATGAGAAAGCGCTGGTATGTTCTTCTGAGGTG 542

QY 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
Db 543 TACCGTAGGGTGGTGTACTATGGAATTTGTTGATTATACAACTATGATGATGAAG 602

QY 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
Db 603 TATGCTATAAGGAAGCTTTGATGATGATCTGATGATTTAAAAATGCCCTTCTCTGAGCGCTAAT 662

QY 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
Db 663 AGGTAAGAGGATGCTGGAAAGAGCGCGCTCTTATTCACGACCGCGTAACAAACC 722

QY 201 ArgSerArgSerArgSerArgGlyProSer 210
Db 723 CGAAGTGGCAGCTACAGCAGGAGTCCAAAT 752

RESULT 9
BO605903
LOCUS
DEFINITION BRY_1489 wheat EST endosperm library Triticum aestivum cDNA 5', mRNA sequence.
ACCESSION BO605903
VERSION BO605903.1 GI:21555035
KEYWORDS EST
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.
1 (bases 1 to 834)
REFERENCE Clarke, B., Lambrecht, M. and Rhee, S.Y.
AUTHORS Arabidopsis genomic information for interpreting wheat EST
TITLE Arabidopsis genomic information for interpreting wheat EST sequences
JOURNAL Funct. Integr. Genomics 3 (1-2), 33-38 (2003)

MEADLINE 22478026
PUBMED 12590341
COMMENT Contact: Lambrecht M
The Arabidopsis Information Resource
Carnegie Institution of Washington, Dept. of Plant Biology
260 Panama Street, Stanford, CA 94305, USA
Tel: 1 650 325 1521 x 251
Fax: 1 650 325 3748
Email: rhee@acoma.stanford.edu.
Location/Qualifiers
1. 834
/organism="Triticum aestivum"
/mol type="mRNA"
/cultivar="Wyuana"
/db_xref="taxon:4565"
/tissue type="endosperm"
/dev stage="developing endosperm tissue 8, 10 and 12 DPA (days post anthesis)"
/clone lib="wheat EST endosperm library"
BASE COUNT 198 a 199 c 234 g 200 t 3 others
ORIGIN

Alignment Scores: 5.84e-26 Length: 834
Pred. No.: 150.00 Matches: 140
Score: 71.43% Conservative: 10
Percent Similarity: 66.67% Mismatches: 60
Best Local Similarity: 67.57% Indels: 0
Query Match: 67.57% Gaps: 0
DB: 13

US-10-014-927-19MOD_COPY_1_222 (1-222) x BO605903 (1-834)

QY 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
Db 123 ATGAGTAGGCGCTGGAGCGGACCATTTAGTTGGGAACCTCCAGGGGATATCAGGAG 182

QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40
Db 183 CGGAGGTGGAGGATCTCTTCAAGATATGAGCGGATAGTTGAAATTCACTTGAAGTC 242

QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspIleAspAsp 60
Db 243 CCCCCAAGGCTCTCGCTTGTGTTGTTGAGTTTGAAGATCCCGCTGATCTGAAGAT 302

QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuValGluIleAla 80
Db 303 GCATTCATGGCCGAGATGGCTACACTGATGGGAATCGGCTTANGTTGAAGTGA 362

QY 81 HisGlyGlyArgArg*****SerTyrSerAlaSer 100
Db 363 CATGGCGGAGGCGCAACTCTTCATCCCTCCGAACAGCTATGTTGGTGGGAGCGCGT 422

QY 101 ArgAlaProSerArgSerArgSerArgValLeuValThrGlyLeuProSerAla 120
Db 423 GGTGGTGTCTTAGGCATACGGAGTATCGTGTCTTCTGTTACGACCTACCTTCTCTGCA 482

QY 121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
Db 483 TCATGGCAAGATCTGAAGGACCATATGAGAAAGCGCTGGTATGTTCTTCTGAGGTG 542

QY 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
Db 543 TACCGTAGGGTGGTGTACTATGGAATTTGTTGATTATACAACTATGATGATGAAG 602

QY 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
Db 603 TATGCTATAAGGAAGCTTTGATGATGATCTGATGATTTAAAAATGCCCTTCTCTGAGCGCTAAT 662

QY 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
Db 663 AGGTAAGAGGATGCTGGAAAGAGCGCGCTCTTATTCACGACCGCGTAACAAACC 722

QY 201 ArgSerArgSerArgSerArgGlyProSer 210
Db 723 CGAAGTGGCAGCTACAGCAGGAGTCCAAAT 752

Unpublished
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3.
Location/Qualifiers
1. .643
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="CPRO12D7"
/tissue type="roots"
/dev stage="in vitro grown stem cuttings"
/lab_host="SOLR"
/clone_lib="potato roots"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University, Tanksley lab;
sequencing: The Institute for Genomic Research. Roots were
isolated from in vitro grown stem cuttings on CM medium.
Roots were isolated two weeks after placing the stem
cuttings from in vitro grown plants on medium."
BASE COUNT 159 a 124 c 167 g 193 t
ORIGIN
Alignment Scores:
Pred. No.: 6,97e-26 Length: 643
Score: 148.00 Matches: 138
Percent Similarity: 74.37% Conservative: 10
Best Local Similarity: 69.35% Mismatches: 51
Query Match: 66.67% Indels: 0
DB: 12 Gaps: 0
US-10-014-927-19MOD_COPY_1_222 (1-222) x BM111827 (1-643)
QY 4 ArgTTPAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluVal 23
DB 47 CGTCTAAGTCGGACTATCATCGTCGGAATCTTCCTGGTGATTCGGGAGAGAAATTA 106
QY 24 GluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProArg 43
DB 107 GAAGATTGTTTACAGTATGTCCTCCATTGGAAATGATTGAACTTCCACTAGA 166
QY 44 ProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaSerIleTyr 63
DB 167 CCACCTGGTTATCGGTTTCGTAGATTGGAATCTCTCGTACTCTGATCATGCCATCCGT 226
QY 64 GlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleHisGlyGly 83
DB 227 GGCGCGTATGGCTATGACTTGTATGGCGCTGCGCTTCCGAGTTGAACTTGCACATGGTGG 286
QY 84 ArgArg*****SeryTyr:SerAlaSerArgAlaPro 103
DB 287 CGAGGATCATCATATATATCCACAGTATGATGAGTAGTGGAGTATCGGTGATTT 346
QY 104 SerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerIle 123
DB 347 TCTAGCGCTCTGACTATCGGTACTGTTCTCTGGACTACCATCTCTGCTTCATGGCAA 406
QY 124 AspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProasp 143
DB 407 GACTTGAAGGATCACATGCGAGCTGGAGATGTCTGCTTCTCAAGTTTTCGAGAT 466
QY 144 ArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspMetLysTyrIleIle 163
DB 467 CGTGACGGATGAGAGGGATTGGGACTATACCACTATGATGATGATGATGATGATGAT 526
QY 164 ArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArg 183

Db 527 AAGAACTTGATGACTCTCTGTTTCGCAATCATCTCTCGAGCATATATTAGGTGGGA 586
QY 184 GluTyrGluSerArgSerValSerArgSerProaspSerLysSeryTyrArgSer 202
DB 587 TAGTATGATAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 643
RESULT 12
CD439565
LOCUS CD439565 852 bp mRNA linear EST 03-JUN-2003
DEFINITION ELO1N0526D06.b Endosperm_5 Zea mays cDNA, mRNA sequence.
ACCESSION CD439565
VERSION CD439565.1 GI:31355208
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 852)
AUTHORS Lai, J., Dey, N., Kim, C.S., Becraft, P., Larkins, B., Linton, E. and
Messing, J.
TITLE Sequencing of the maize endosperm ESTs
JOURNAL Unpublished
COMMENT Contact: Lai, Jinheng
Dr. Joachim Messing's lab
Waksman Institute, Rutgers University
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
Tel: 732-445-3801
Fax: 732-445-5735
Email: jlai@waksman.rutgers.edu
Seq primer: T3.
Location/Qualifiers
1. .852
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W22"
/db_xref="taxon:4577"
/tissue type="Endosperm of 7-23DAP"
/clone_lib="Endosperm_5"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI"
BASE COUNT 217 a 200 c 225 g 210 t
ORIGIN
Alignment Scores:
Pred. No.: 4.11e-25 Length: 852
Score: 147.00 Matches: 137
Percent Similarity: 68.69% Conservative: 10
Best Local Similarity: 64.02% Mismatches: 67
Query Match: 66.22% Indels: 0
DB: 14 Gaps: 0
US-10-014-927-19MOD_COPY_1_222 (1-222) x CD439565 (1-852)
QY 1 MetSerSerArgTTPAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
DB 75 ATGACCGGCGGAGCGCTCTAGTACTAGTGGGACCTCCCGCGACATCCGCGAG 134
QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40
DB 135 AGGGAAGTGGATGATCTCTTCTACAGTATGGAGTATAGTGGAAATTGACTTGAAAT 194
QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60
DB 195 CCACCAAGGCTCTCGGTTTGTGCTTTTGTGTTGAGTTGAGGACGACCATGATGCTGAAG 254
QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
DB 255 GCATATATGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 314
QY 81 HisGlyGlyArgArg*****SeryTyr:SerAlaSer 100
DB 315 CATGCTGAGCGGACATCTCTTTTGTATCGATCTAGCAGCTATAGCAGTCTGCTGGACAA 374

QY	101	ArgAlaProSerArgSerArgSerArgValLeuValThrGlyLeuProProSerAla	120
Db	375	CGCGTGCTCAAACGTTCTGATACCGTGTATGTTACTGATTAACCTTCTAGCA	434
QY	121	SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal	140
Db	435	TCATGGCAAGATCTCAAGGACCATATCGCGGAGCTGGTGATGTCGTCTTTCACTGATGTG	494
QY	141	PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys	160
Db	495	TATCGTAGGCTGGAGCAACTATTGGATAGTCTGATTATATACTAATAAGATATGAAA	554
QY	161	TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle	180
Db	555	CACGCGAAGAGAGCTAGATGATCTGAGTCCGTAATGCTTTTTCAGGACATATGTC	614
QY	181	ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr	200
Db	615	CGGTGAGGAGTATGATGCTAGCGGAGCGGCTTCGCTCCAGAGCGAACCCTCT	674
QY	201	ArgSerArgSerArgSerArgGlyProSerCysSerTyrSer	214
Db	675	AAGTCAAGAGAGAGGACAGGACAGGACCTCGTCAAGAGC	716
RESULT 13			
AY108591			
LOCUS	AY108591	1255 bp mRNA linear HTC 17-OCT-2002	
DEFINITION	Zea mays PCO125156 mRNA sequence.		
ACCESSION	AY108591		
VERSION	AY108591.1 GI:21211708		
KEYWORDS	HTC.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD		
TITLE	clade; Panicoidae; Andropogoneae; Zea.		
JOURNAL	1 (bases 1 to 1255)		
REFERENCE	Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,		
AUTHORS	Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.		
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of		
JOURNAL	Overgo Probes		
REFERENCE	Unpublished (2002)		
AUTHORS	2 (bases 1 to 1255)		
TITLE	Coe,E.H.		
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of		
COMMENT	Missouri, Columbia, MO 65211, USA		
	If you are interested in getting corresponding physical clones,		
	these are publicly available from ZmDB and may be found by BLAST		
	searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,		
	www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the		
	maize cDNA sequences is either Virginia Walbot, Stanford or Pat		
	Schnable, Iowa State, then clones may be requested from ZmDB:		
	www.zmdb.iastate.edu.		
FEATURES	Location/Qualifiers		
source	1..1255		
	/organism="Zea mays"		
	/mol_type="mRNA"		
	/db_xref="maizedb:637293"		
	/db_xref="taxon:4577"		
	/clone_lib="Maize Mapping Project/DuPont Consensus		
	Library"		
	/note="this sequence is part of a project of EST		
	assemblies resulting from the application of public		
	configs to seed Dupont contigs; this resource was		
	assembled by Dupont as part of a collaboration for the		
	overgo addressing of BACs in conjunction with the Maize		
	Mapping Project"		
BASE COUNT	337 a 287 c 315 g 316 t		
ORIGIN			
Alignment Scores:			

Pred. No.:	2,746-24	Length:	1255
Score:	146.50	Matches:	148
Percent Similarity:	71.17%	Conservative:	10
Best Local Similarity:	66.67%	Mismatches:	61
Query Match:	65.99%	Indels:	3
DB:	11	Gaps:	1
US-10-014-927-19MOD_COPY_1_222 (1-222) x AY108591 (1-1255)			
QY	1	MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys	20
Db	126	ATGAGCAGCGCTGGAGCGGCACGATCTACGTCGGGAACCTCCCGCGGACATCCGGAG	185
QY	21	CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle	40
Db	186	AGGAGGTGGAGATCTGTCTTACAGATGATGTAATAATTGTGACATTGACCTGAGGTC	245
QY	41	ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp	60
Db	246	CCCCCAAGACCACTGGTTATGCTTTTGTGAGTTTGAAGATCCTCTCGTGATGCTGAGG	305
QY	61	AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla	80
Db	306	GCAATTGCTGGACGGATGGATACAACTTGTATGACACACCTCTAAGAGTGGAGCTGCT	365
QY	81	HisGlyGlyArg-----Arg*****Arg*****SerTyr	97
Db	366	CATGTGTGTAGAGGTAATGCTTCTCGCATGCTTCAAGTGGCTTTGGTGGCGTGGT	425
QY	98	SerAlaSerArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuPro	117
Db	426	GGAGCAGCTCGTGGTGTGTCGAGACACTCAGAGTATCGTGTCTTGTCTGATGCTGCT	485
QY	118	ProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPhe	137
Db	486	TCTTCTGCATCATGCGAGGATTTAAAGGATCATATGCGGAAGCTGGTGTGTTTGTTC	545
QY	138	SerGluValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAsp	157
Db	546	TCTGAAGTGTATCGGGAAGCGCGTGGACCGCTAGGATTTGTGGACTACACAAATATAT	605
QY	158	AspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSer	177
Db	606	GATATGAATATGCTATAAAGAGCTGGATGATGATCTGAATTCAGGAAGCCCTTTGGG	665
QY	178	AlaTyrIleArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSer	197
Db	666	GCCTATATAGGTGTAGGAATATAACGGCAACCTGGTGGCTCTACTCAGTAGCCGA	725
QY	198	LysSerTyrArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSer	217
Db	726	AGCCCAAGTCGTAGTTACAGCAAAAGCAGGAGTCCGAGTAAATCACCAGGACTCGCGT	785
QY	218	ArgSer 219	
Db	786	TCATCA 791	
RESULT 14			
LOCUS	CA008486	614 bp mRNA linear EST 23-OCT-2002	
DEFINITION	HUI1B23r HU Hordeum vulgare subsp. vulgare cDNA clone HUI1B23		
	5-PRIME, mRNA sequence.		
ACCESSION	CA008486		
VERSION	CA008486.1 GI:24285468		
KEYWORDS	EST.		
SOURCE	Hordeum vulgare subsp. vulgare		
ORGANISM	Hordeum vulgare subsp. vulgare		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae		
	; Triticeae; Hordeum.		
	1 (bases 1 to 614)		
REFERENCE	Zhang,H., Potokina,E., Michalek,W., Weschke,W., Stein,N. and Graner		
AUTHORS	,A.		

TITLE
JOURNAL
COMMENT

Barley ESTs from germinating seeds
Unpublished
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 614 Std Error: 0.00
Plate: 11 row: B column: 23
Seq primer: M13rev.

FEATURES
Location/Qualifiers
1..614
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/db_xref="GABI:266136"
/db_xref="taxon:112509"
/clone="HUI1B23"
/tissue_type="germinating seeds"
/dev_stage="germinating seeds, 16-48 h"
/lab_host="XL10-Gold"
/clone_lib="HU"
/note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of cDNA); Site 2: XhoI (3'-end of cDNA); germinating seeds, 16-48h. Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also due to the cloning system used Blue/white selection for recombinants is not 100% reliable. Average insert size is 1.2 kb"

BASE COUNT 146 a 118 c 189 g 161 t

ORIGIN
source

Alignment Scores:
Pred. No.: 28-25 Length: 614
Score: 146.00 Matches: 136
Percent Similarity: 77.25% Conservative: 10
Best Local Similarity: 71.96% Mismatches: 43
Query Match: 65.77% Indels: 0
DB: 14 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x CA008486 (1-614)

QY 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
Db 48 ATGAGTAGGCGCTGGAGCGGACCATTTACGTTGGGAACCTCCAGGGGATATCAGGGAG 107

QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLys 40
Db 108 CGGGAGGTGGAGGATCTGTTCACAGATATGACCGGATAGTGGAAATGACTTGAAGTTC 167

QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
Db 168 CCCCCAGCCACCTGGCTTCTGTTTGTAGTTTGAGATCCCGGTGATGCCGAGAT 227

QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
Db 228 GCGATTACGGCGGAGATGCTTACAACTTTGATGGAATCGCTGAGGTTGAACTTCCA 287

QY 81 HisGlyGlyArgArg*****SerTyrSerAlaSer 100
Db 288 CATGTGGGAGGCGCTAACTCTTCATCCCTTCTAACAGCCATGTGGTGGAGAGCGCGT 347

QY 101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla 120
Db 348 GTGGGTCTCTAGGCATACAGAGTATCGTTCCTGTTTACGTGAGTACCTCTTCTTCTCA 407

QY 121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
Db 408 TCATGCCAGATCTAAAGGATCATATGAGAAAGGCTGGTGTATGTTTCTCTGAGGTG 467

QY 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
Db 468 TATCGTGAGGCGTGGTACTATTGGAATTTGATATACAACTATGATGATATGAG 527

QY 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
Db 528 TATGCTATAGGAAGCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 507

QY 181 ArgValArgGluTyrGluSerArgSer 189
Db 588 AGGCTGAAGGAGTATGCTGGCAAAAGC 614

RESULT 15
BU269490 625 bp mRNA linear EST 09-APR-2002
LOCUS BU269490 Y. Ogihara unpublished cDNA library, Wh_oh Triticum
DEFINITION aestivum cDNA clone whoh22d09 5', mRNA sequence.
ACCESSION BU269490
VERSION BU269490.1 GI:20095590
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 625)
AUTHORS Ogihara, Y. and Murai, K.
TITLE Expressed genes in Triticum aestivum
JOURNAL Unpublished
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..625
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whoh22d09"
/tissue_type="pistil at heading date"
/dev_stage="Peak", scale 10.5"
/clone_lib="Y. Ogihara unpublished cDNA library, Wh_oh"

BASE COUNT 148 a 126 c 190 g 161 t

ORIGIN
source

Alignment Scores:
Pred. No.: 2.15e-25 Length: 625
Score: 146.00 Matches: 136
Percent Similarity: 77.25% Conservative: 10
Best Local Similarity: 71.96% Mismatches: 43
Query Match: 65.77% Indels: 0
DB: 12 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x BU269490 (1-625)

QY 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
Db 54 ATGAGTAGGCGCTGGAGCGGACCATTTACGTTGGGAACCTCCAGGGGATATCAGGGAG 113

QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLys 40
Db 114 CGGGAGGTGGAGGATCTCTTCTACAAATGATGGAGGATAGTGGAAATGACTTGAAGTTC 173

QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
Db 174 CCCCCAGGCGCTCTGGCTTGTGTTTGTAGTTTGAAGATCCCGGTGATGCTGAAGAT 233

QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80

Search completed: February 4, 2004, 22:08:06
Job time : 2049 secs.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 15:22:45 ; Search time 14428 Seconds
(without alignments)
11466.491 Million cell updates/sec

Title: US-10-014-927-18

Perfect score: 4044

Sequence: 1 agacaagatgcttactctt.....atgattatcttaactaac 4044

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045491386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_on.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_to.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_nam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	4044	100.0	4044	6	AX040661	AX040661 Sequence
2	4021	99.4	5164	8	ATH131214	ATH131214 Arabidops
3	3493.2	86.4	59261	8	T12M4	AC003114 Arabidops
4	780	19.3	780	6	AX509198	AX509198 Sequence
5	571	14.1	110514	8	F7G19	AC000106 Sequence
6	311.4	7.7	933	8	AY050912	AY050912 Arabidops
7	139	3.4	109465	8	AP06409	AP06409 Lotus jap
8	131.4	3.2	133656	2	AC144541	AC144541 Medicago
9	117	2.9	838	8	AY150486	AY150486 Arabidops
10	113.8	2.8	939	8	AY128356	AY128356 Arabidops
11	108.8	2.7	762	6	AX506504	AX506504 Sequence
12	107.6	2.7	120374	8	AC126791	AC126791 Medicago
13	107.6	2.7	129199	2	AC134522	AC134522 Medicago
14	105.4	2.6	80367	8	AC012329	AC012329 Arabidops
15	105.4	2.6	104204	8	ATT9C5	AL132964 Arabidops
16	98.2	2.4	107367	2	AC136288	AC136288 Medicago
17	86.6	2.1	168267	8	AP003535	AP003535 Oryza sat
18	78.6	1.9	5285	6	AR204400	AR204400 Sequence
19	78.6	1.9	94536	8	T14P8	AF069298 Arabidops
20	78.6	1.9	194892	8	ATCHRIV6	AL161494 Arabidops
21	78	1.9	157248	8	AP004300	AP004300 Oryza sat
22	77.8	1.9	93791	2	AC138073	AC138073 Homo sapi
23	75.8	1.9	7218	6	I66494	I66494 Sequence 14
24	75	1.9	66993	2	AC138074	AC138074 Homo sapi
25	70.4	1.7	761	11	CNS06KRG	AL402998 T7 end of
26	67.4	1.7	664	6	E23356	E23356 Virus vecto
27	67.4	1.7	7372	6	E23357	E23357 Virus vecto
28	67.4	1.7	7797	6	E23355	E23355 Virus vecto
29	67.4	1.7	7996	6	E23359	E23359 Virus vecto
30	67.4	1.7	80216	2	AC139363	AC139363 Homo sapi
31	67	1.7	289973	2	AC135678	AC135678 Rattus no
32	65.4	1.6	4146	8	AF001035	AF001035 Arabidops
33	65.4	1.6	6161	6	AX345313	AX345313 Sequence
34	65.4	1.6	100239	8	AC009525	AC009525 Arabidops
35	65.2	1.6	286208	2	AC117140	AC117140 Rattus no
36	64.8	1.6	14006	6	AX346860	AX346860 Sequence
37	64.8	1.6	93975	2	AC139011	AC139011 Homo sapi
38	64.4	1.6	6171	6	AX345690	AX345690 Sequence
39	64.2	1.6	127709	2	CNS05TCL	AL352983 Homo sapi
40	64.2	1.6	156550	2	AC015830	AC015830 Homo sapi
41	64	1.6	151368	2	AC142664	AC142664 Macaca mu
42	63.8	1.6	143585	2	AC013349	AC013349 Homo sapi
43	63.2	1.6	6668	6	AX346599	AX346599 Sequence
44	63	1.6	83440	2	AC024285	AC024285 Homo sapi
45	62.8	1.6	65034	2	AC016067	AC016067 Homo sapi

ALIGNMENTS

RESULT 1
AX040661
LOCUS AX040661 4044 bp DNA linear PAT 23-NOV-2000
DEFINITION Sequence 18 from Patent WO0065059.
ACCESSION AX040661
VERSION AX040661.1 GI:11340389
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.
REFERENCE
1 Barta,A., Lopato,S., Kalyana.M. and Dörner,S.
AUTHORS
TITLE Splicing factor
JOURNAL Patent: WO 0065059-A 18 02-NOV-2000;
\sterreichisches Forschungszentrum Seibersdorf Ges.m.b.H; . (AT)

FEATURES		Location/Qualifiers			
source		1. .4044			
		/organism="unidentified"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:32644"			
		/note="Beschreibung des unbekannten ORGANISM: Genom atSRp30"			
BASE COUNT		1061 a	727 c	829 g	1427 t
ORIGIN					
Query Match		100.0%;	Score 4044;	DB 6;	Length 4044;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 4044;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	AGACAAAGATGCTTACTTCTTAAACATGTTTCGAGGTTTATTGAAATGATCACAGCTTC	60		
Db	1	AGACAAAGATGCTTACTTCTTAAACATGTTTCGAGGTTTATTGAAATGATCACAGCTTC	60		
QY	61	TAACTATGGTATCTTCTTCCCTGCAAGCGAACAGTGGGAAGATGATTGATAACGGAATAT	120		
Db	61	TAACTATGGTATCTTCTTCCCTGCAAGCGAACAGTGGGAAGATGATTGATAACGGAATAT	120		
QY	121	CGGAACATCATCTCAACAAACCAAAAAATTGGACATCATATCGCAACAAATTCATATAGGAA	180		
Db	121	CGGAACATCATCTCAACAAACCAAAAAATTGGACATCATATCGCAACAAATTCATATAGGAA	180		
QY	181	AAATACTGAAATTCCTAAACAGAAAAACCAACCGAACAGACGAGCACTCAACGACTGA	240		
Db	181	AAATACTGAAATTCCTAAACAGAAAAACCAACCGAACAGACGAGCACTCAACGACTGA	240		
QY	241	GAGAGACCGTGGACGGTGTCTACGGCGGAAAAATGGTCTTGGACGGAGTTACTAATCGGCGA	300		
Db	241	GAGAGACCGTGGACGGTGTCTACGGCGGAAAAATGGTCTTGGACGGAGTTACTAATCGGCGA	300		
QY	301	ATTGAGATTTGAGAGGTGGTAGTAGAGAAACCGAGAGAAATGTTTCTCTCAAAAAATCCC	360		
Db	301	ATTGAGATTTGAGAGGTGGTAGTAGAGAAACCGAGAGAAATGTTTCTCTCAAAAAATCCC	360		
QY	361	CAAGTGTTCGGATCTAGTGTCTCTTTGTCGAAAAACGACAGTGTGTAGGAAACCTAGG	420		
Db	361	CAAGTGTTCGGATCTAGTGTCTCTTTGTCGAAAAACGACAGTGTGTAGGAAACCTAGG	420		
QY	421	AGAATGAATGACCGATGGTCCGAATCCGATTCGAAATTTGGTTCGAATTTGTAATACTAA	480		
Db	421	AGAATGAATGACCGATGGTCCGAATCCGATTCGAAATTTGGTTCGAATTTGTAATACTAA	480		
QY	481	CATACAAATATTCGGTTTGAATGATAAGAAAAACACATTCGATCCGGTTAGAACAAATAT	540		
Db	481	CATACAAATATTCGGTTTGAATGATAAGAAAAACACATTCGATCCGGTTAGAACAAATAT	540		
QY	541	ATTAAACAGCCCATTAACAAATATGGCGCGATCTTGATCAACTGGGCTATTTCATCGTTGA	600		
Db	541	ATTAAACAGCCCATTAACAAATATGGCGCGATCTTGATCAACTGGGCTATTTCATCGTTGA	600		
QY	601	TACATGGCGCGCACAGGATTAATCCAGTTCGGTTTATTAAGAGGATAGTATTTCCA	660		
Db	601	TACATGGCGCGCACAGGATTAATCCAGTTCGGTTTATTAAGAGGATAGTATTTCCA	660		
QY	661	AACGAAACGGTGGTGTCTCTTCCAGAACAACTTAACGGTTTCTCGAAACATCTTCTTCT	720		
Db	661	AACGAAACGGTGGTGTCTCTTCCAGAACAACTTAACGGTTTCTCGAAACATCTTCTTCT	720		
QY	721	TCCTCTTCTCGAAATATTTTCCAGTAATCAATTTCTCTCTCTCTCTCTCTCTCTCTCTCT	780		
Db	721	TCCTCTTCTCGAAATATTTTCCAGTAATCAATTTCTCTCTCTCTCTCTCTCTCTCTCTCT	780		
QY	781	AACATAATTTCTGCTCTGAGGTATCAGATGATGAGCGGATGGAATTCGTAGCTCTAGTT	840		
Db	781	AACATAATTTCTGCTCTGAGGTATCAGATGATGAGCGGATGGAATTCGTAGCTCTAGTT	840		
QY	841	GGGAATTTCCCTGGAGATATTTCGCAAGTGTGAGGTTGAAGATCTCTTCTCAAGAGTTTGA	900		
QY	841	GGGAATTTCCCTGGAGATATTTCGCAAGTGTGAGGTTGAAGATCTCTTCTCAAGAGTTTGA	900		

Qy	901	AAATTTCTCTCTCTCTCGATAAAAAATGAATTCATTATGACTAGTTTGGGTTTCATAAA	960
Db	901	AAATTTCTCTCTCTCTCGATAAAAAATGAATTCATTATGACTAGTTTGGGTTTCATAAA	960
Qy	961	TTTGCAATCTGCTCTGCTGAGACAAATTTAAATCGACTCTTATGATATTTGTTTCAGTA	1020
Db	961	TTTGCAATCTGCTCTGCTGAGACAAATTTAAATCGACTCTTATGATATTTGTTTCAGTA	1020
Qy	1021	TGGACCAATTTGGACATTTGATTTGAAGATTTCCACCGAGACCTCTCGTTATGCTTTGT	1080
Db	1021	TGGACCAATTTGGACATTTGATTTGAAGATTTCCACCGAGACCTCTCGTTATGCTTTGT	1080
Qy	1081	CGAGCTATATTGATCAAGTACAAATTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCT	1140
Db	1081	CGAGCTATATTGATCAAGTACAAATTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCT	1140
Qy	1141	ATGACTAAGATAGTTTGTATTTGGTGGCAGTTTGAAGATTCCTGCTGATGCGACGATGCA	1200
Db	1141	ATGACTAAGATAGTTTGTATTTGGTGGCAGTTTGAAGATTCCTGCTGATGCGACGATGCA	1200
Qy	1201	ATTTATGACGCTGATGTTATGATTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1260
Db	1201	ATTTATGACGCTGATGTTATGATTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1260
Qy	1261	ATGAAAGCTAGCTTAATTTCTGTAATTTCTTTGTAAGGTTGTTACTTTGTTGATGTTT	1320
Db	1261	ATGAAAGCTAGCTTAATTTCTGTAATTTCTTTGTAAGGTTGTTACTTTGTTGATGTTT	1320
Qy	1321	TTAGGTTGAGATTTGCATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1380
Db	1321	TTAGGTTGAGATTTGCATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1380
Qy	1381	CAGCTACAGTGGACCGCTGCACCTTCAAGACGCTCTGACTACCGCGTTTGTAGAGTCT	1440
Db	1381	CAGCTACAGTGGACCGCTGCACCTTCAAGACGCTCTGACTACCGCGTTTGTAGAGTCT	1440
Qy	1441	TCCTGATTTGTTATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1500
Db	1441	TCCTGATTTGTTATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1500
Qy	1501	CTAAACATGTTAGTGTCTGTGACCGGATTTACCGCTTCTGCTTCTGCGGACGACCTTAA	1560
Db	1501	CTAAACATGTTAGTGTCTGTGACCGGATTTACCGCTTCTGCTTCTGCGGACGACCTTAA	1560
Qy	1561	TAAGGACACATATATGATGCTTTTCTCTGAATGTTGGTCTCTATATCATGTTTGGATT	1620
Db	1561	TAAGGACACATATATGATGCTTTTCTCTGAATGTTGGTCTCTATATCATGTTTGGATT	1620
Qy	1621	TATCTCTTTCTGAAATGATGTTATTTGCTATTTACGGTGTATTTAGGATCATGCGCAA	1680
Db	1621	TATCTCTTTCTGAAATGATGTTATTTGCTATTTACGGTGTATTTAGGATCATGCGCAA	1680
Qy	1681	AGCTCGAGATGCTCTCTCTGAAATTTCCCTGACCGTAAAGGTGAGTTGACATTCGA	1740
Db	1681	AGCTCGAGATGCTCTCTCTGAAATTTCCCTGACCGTAAAGGTGAGTTGACATTCGA	1740
Qy	1741	TAGTTTGGATAGCTTTTGTATTTGATGTTAGTAAATTTAGTCTTTGTGAAGGAAATAG	1800
Db	1741	TAGTTTGGATAGCTTTTGTATTTGATGTTAGTAAATTTAGTCTTTGTGAAGGAAATAG	1800
Qy	1801	GTGTTAAGCATCTGAACCTGCTAAACATCATTCAGTATTTCTTTGTAGGCAATGCTGGG	1860
Db	1801	GTGTTAAGCATCTGAACCTGCTAAACATCATTCAGTATTTCTTTGTAGGCAATGCTGGG	1860
Qy	1861	TTGTCGATTTATAGCAACTATGATGATGAAAGTACGCAAGTAAATTTTATATCTTTGCA	1920
Db	1861	TTGTCGATTTATAGCAACTATGATGATGAAAGTACGCAAGTAAATTTTATATCTTTGCA	1920
Qy	1921	GCAATGTTCTCGACTTATGCTTTAGACTGCTTTGTTTCTATAGTATACCGAGCTGAAT	1980
Db	1921	GCAATGTTCTCGACTTATGCTTTAGACTGCTTTGTTTCTATAGTATACCGAGCTGAAT	1980

QY	1981	TTATCTTCTCGAGGCCAGTGTGGATCTTTGATGTTCCCTTAAATTTTGGATGACA	2040
DB	1981	TTATCTTCTCGAGGCCAGTGTGGATCTTTGATGTTCCCTTAAATTTTGGATGACA	2040
QY	2041	GATAGGAAACCTTGATCCACCTGAATTCGAAATGCTTCTAGTGTCTTATATACGGGT	2100
DB	2041	GATAGGAAACCTTGATCCACCTGAATTCGAAATGCTTCTAGTGTCTTATATACGGGT	2100
QY	2101	ATGTTGTAATGCTTCTTCTTTGATTTTGTAAAGCATAAAGTGGATATGAGTCACTCTGAAT	2160
DB	2101	ATGTTGTAATGCTTCTTCTTTGATTTTGTAAAGCATAAAGTGGATATGAGTCACTCTGAAT	2160
QY	2161	TTACTGTTTCAGGTGAGGGAATATGAGTCGAGGAGTGTGAGTCGAAGCCACAGATGTTCTA	2220
DB	2161	TTACTGTTTCAGGTGAGGGAATATGAGTCGAGGAGTGTGAGTCGAAGCCACAGATGTTCTA	2220
QY	2221	AAAGCTATAGAGAGGAGTCGAGCCGTGGTCCAAAGCTGTAGCTATATAGTACGAAGACA	2280
DB	2221	AAAGCTATAGAGAGGAGTCGAGCCGTGGTCCAAAGCTGTAGCTATATAGTACGAAGACA	2280
QY	2281	GGAGGCTTTTTTTTTTTTTTTTTTTTCAATAAACCTTAAGACATATAAGGGATTTTATTTGT	2340
DB	2281	GGAGGCTTTTTTTTTTTTTTTTTTTTCAATAAACCTTAAGACATATAAGGGATTTTATTTGT	2340
QY	2341	AACTTATTTATGAAATTAACCTCACTTCTAAATGCAATGCAGTGTCTCACCTGCTAGATCC	2400
DB	2341	AACTTATTTATGAAATTAACCTCACTTCTAAATGCAATGCAGTGTCTCACCTGCTAGATCC	2400
QY	2401	ATTTCCCGCGTTCACGGCCCTTCTAGTCTTCTCGCTCGCTATACAGTCTGTCTCAAGG	2460
DB	2401	ATTTCCCGCGTTCACGGCCCTTCTAGTCTTCTCGCTCGCTATACAGTCTGTCTCAAGG	2460
QY	2461	TATGAGTGTATGATTTGAT	2520
DB	2461	TATGAGTGTATGATTTGAT	2520
QY	2521	TCTTGCATAGTGAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2580
DB	2521	TCTTGCATAGTGAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2580
QY	2581	CGGTACACAGTGAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	2640
DB	2581	CGGTACACAGTGAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	2640
QY	2641	CTTGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2700
DB	2641	CTTGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2700
QY	2701	TTACCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2760
DB	2701	TTACCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2760
QY	2761	TTTTTTATGTTTGTAGTCTTGCACAGAGATGCTCCCATCATATTTAGTCCCTTTCCCTTTC	2820
DB	2761	TTTTTTATGTTTGTAGTCTTGCACAGAGATGCTCCCATCATATTTAGTCCCTTTCCCTTTC	2820
QY	2821	TCCTTGTGTCGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2880
DB	2821	TCCTTGTGTCGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2880
QY	2881	TCCAGCGAGGTGAATTTATACGATCAACCTCAATATCCGACCTATAATTTAAAGAAC	2940
DB	2881	TCCAGCGAGGTGAATTTATACGATCAACCTCAATATCCGACCTATAATTTAAAGAAC	2940
QY	2941	ACTTAGCTAGATGTTTCACTTTTGAATAATTTATTTCTATCGGAAGGGGCGAGATGTTCT	3000
DB	2941	ACTTAGCTAGATGTTTCACTTTTGAATAATTTATTTCTATCGGAAGGGGCGAGATGTTCT	3000
QY	3001	GAGGATGCTTCGATCAATTTACCGTGATCAATTTGTTGCGTTGTTGTTGTTGTTGTTGTTGTTG	3060
DB	3001	GAGGATGCTTCGATCAATTTACCGTGATCAATTTGTTGCGTTGTTGTTGTTGTTGTTGTTG	3060
QY	3061	CTGGCTCACTGCTACGAGCTGGGATTTGATCTAGATGGTCACTAGATGGATCTTTGG	3120

DB	3061	CTGGCTCACTGCTACGAGCTGGGATTTGGATCTAGATGGTCACTAGATGGATCTTTGG	3120
QY	3121	ACTGGATTTACAAAGCTGGATTAAGCACTGAACTTCTGTTTACGGTCTGGTCTGGT	3180
DB	3121	ACTGGATTTACAAAGCTGGATTAAGCACTGAACTTCTGTTTACGGTCTGGTCTGGT	3180
QY	3181	CTGCTACTCCGCGCTATCAGCTGTAGGATCTGATCGAAAGTTTGGACTATGATTACT	3240
DB	3181	CTGCTACTCCGCGCTATCAGCTGTAGGATCTGATCGAAAGTTTGGACTATGATTACT	3240
QY	3241	CTGATTCCTCAATATATTTATCTTTTTCACATAGTGGATTTCTGTTGTTGTTTCTTCT	3300
DB	3241	CTGATTCCTCAATATATTTATCTTTTTCACATAGTGGATTTCTGTTGTTGTTTCTTCT	3300
QY	3301	AGGACAGCATTTAAGCTCCCGGACTAGATGGGAGATGGTCACTAAATTTCTTTGTTATG	3360
DB	3301	AGGACAGCATTTAAGCTCCCGGACTAGATGGGAGATGGTCACTAAATTTCTTTGTTATG	3360
QY	3361	CCACCTTACATGGGTTTTCGGTCTTCTGCTGAGGTCCTCAATCAAGATCAAAATCAAGA	3420
DB	3361	CCACCTTACATGGGTTTTCGGTCTTCTGCTGAGGTCCTCAATCAAGATCAAAATCAAGA	3420
QY	3421	TCAAGATCAAGATCGAATTTCCAGTTTCCAGTGTGTTAGTCTAAAGCTGAACTTCT	3480
DB	3421	TCAAGATCAAGATCGAATTTCCAGTTTCCAGTGTGTTAGTCTAAAGCTGAACTTCT	3480
QY	3481	TTAATTCACAAATCCATGTTTGTTTTAAATACCTGCTCACTTTGTTGTTTCTTCAATCAA	3540
DB	3481	TTAATTCACAAATCCATGTTTGTTTTAAATACCTGCTCACTTTGTTGTTTCTTCAATCAA	3540
QY	3541	CACCACTTAAAGAAATCATGAGACAGACATATAAAATTTGAAGAGTCTGTAGAACGACTA	3600
DB	3541	CACCACTTAAAGAAATCATGAGACAGACATATAAAATTTGAAGAGTCTGTAGAACGACTA	3600
QY	3601	GGTCTACCAACCTCTGTTGTCACATAAATCGCTCTCCAGTGTTCAGCAACATAAT	3660
DB	3601	GGTCTACCAACCTCTGTTGTCACATAAATCGCTCTCCAGTGTTCAGCAACATAAT	3660
QY	3661	CTACCTCTGCTGTTGTTATCATTTCTTCTTAAACGGTATTACATATTATGTTTGG	3720
DB	3661	CTACCTCTGCTGTTGTTATCATTTCTTCTTAAACGGTATTACATATTATGTTTGG	3720
QY	3721	CAGGTGATATCTGGTTGAAATGAAACCTGGCACTGGCTGTACCCGAATGCTCTCAAGC	3780
DB	3721	CAGGTGATATCTGGTTGAAATGAAACCTGGCACTGGCTGTACCCGAATGCTCTCAAGC	3780
QY	3781	TTCTCAGGCTCCACTGCTTAATAGAAATTTGATTCGGAATTTGGGATTTATTACTGGTCTC	3840
DB	3781	TTCTCAGGCTCCACTGCTTAATAGAAATTTGATTCGGAATTTGGGATTTATTACTGGTCTC	3840
QY	3841	TTGTATGGGACGACCAATATGTTCTTCTAGTTTGTAGTGTGACCTGGAAATGCTGTT	3900
DB	3841	TTGTATGGGACGACCAATATGTTCTTCTAGTTTGTAGTGTGACCTGGAAATGCTGTT	3900
QY	3901	ATTGTGTCAATTAAGCGGAACTCTGCTCGGCTGCATATAAAGTTTCATCAGACAT	3960
DB	3901	ATTGTGTCAATTAAGCGGAACTCTGCTCGGCTGCATATAAAGTTTCATCAGACAT	3960
QY	3961	TGTTTGGGTTGGTGGGTTTTTCCATACATATACATTTACATTAACAATCTGTTGTC	4020
DB	3961	TGTTTGGGTTGGTGGGTTTTTCCATACATATACATTTACATTAACAATCTGTTGTC	4020
QY	4021	TTTTTATGATTTATCTTAAACTAAAC 4044	
DB	4021	TTTTTATGATTTATCTTAAACTAAAC 4044	

RESULT 2
 ATH131214
 LOCUS 5164 bp DNA linear PLN 07-MAY-1999
 DEFINITION Arabidopsis thaliana srp30 gene, exons 1-12.
 ACCESSION AJ131214

AJ131214.1 GI:4775269
SR2/ASF-like splicing modulator; srp30 gene.
Arabidopsis thaliana (thale cress)
ORIGIN
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
Lopato,S., Kalyana,M., Dörner,S., Kobayashi,R., Krainer,A.R. and
Barta,A.
TITLE
atsRp30, one of two SR2/ASF-like proteins from Arabidopsis
thaliana, regulates splicing of specific plant genes
JOURNAL
Genes Dev. 13 (8), 987-1001 (1999)
MEDLINE
99234087
PUBMED
10215626
REFERENCE
2 (bases 1 to 5164)
Barta,A.
Direct Submission
Submitted (02-DEC-1998) Barta A., Institute for Biochemistry,
University of Vienna, Dr. Bohrgasse 9/3, A-1030 Vienna, AUSTRIA
FEATURES
source
1. .5164
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
join(1806..2014,2138..2203,2290..2364,2444..2546,
2632..2678,2786..2843,2968..3017,3161..3217,3291..3403,
3501..3578,4515..4572,4844..5125)
1806..2014
/number=1
1928..4858
/gene="srp30"
join(1928..2014,2138..2203,2290..2364,2444..2546,
2632..2678,2786..2843,2968..3017,3161..3217,3291..3403,
3501..3578,4515..4572,4844..4858)
/gene="srp30"
/codon_start=1
/product="SF2/ASF-like splicing modulator Srp30"
/protein_id="CA842557.1"
/db_xref="GI:4775270"
translations="MSRWRTIYVGNLPGDIRKCEVDLFYKYGPIVDILKIPRP
PGYAFVEFPRADDAIYGRDGYDFDGCRLVEIAHGGRSPSPVDYSSYSASRA
PSRSYRVLTGLPPSASQDLKDMRKAGDVCSEVPDRKMGSGVDYSNYDDMK
YAIKLDATFRNFAFSAYIRVREYSRSVSRSPDSDSKYRSRSGRSPSCYSKSR
SVSPARISIPRSPPLSRSLYSVSRSGSRSRSGRSPSPVPIG"
join(1928..2014,2138..2203,2290..2364,2444..2546,
2632..2678,2786..2843,2968..3017,3161..3217,3291..3403,
3501..3578,4178..4214)
/gene="srp30"
/note="alternative exon 10a"
/codon_start=1
/product="SF2/ASF-like splicing modulator Srp30, variant
1"
/protein_id="CA842558.1"
/db_xref="GI:4775271"
translations="MSRWRTIYVGNLPGDIRKCEVDLFYKYGPIVDILKIPRP
PGYAFVEFPRADDAIYGRDGYDFDGCRLVEIAHGGRSPSPVDYSSYSASRA
PSRSYRVLTGLPPSASQDLKDMRKAGDVCSEVPDRKMGSGVDYSNYDDMK
YAIKLDATFRNFAFSAYIRVREYSRSVSRSPDSDSKYRSRSGRSPSCYSKSR
SVSPARISIPRSPPLSRSLYSVSRSGSRSLRAGDWI"
2015..2137
/gene="srp30"
/number=1
2138..2203
/gene="srp30"
/number=2
2204..2289
/gene="srp30"
/number=2
2290..2364
/gene="srp30"
/number=3
2365..2443

/gene="srp30"
/number=3
2444..2546
/gene="srp30"
/number=4
2547..2631
/gene="srp30"
/number=4
2632..2678
/gene="srp30"
/number=5
2679..2785
/gene="srp30"
/number=5
2786..2843
/gene="srp30"
/number=6
2844..2967
/gene="srp30"
/number=6
2968..3017
/gene="srp30"
/number=7
3018..3160
/gene="srp30"
/number=7
3161..3217
/gene="srp30"
/number=8
3218..3290
/gene="srp30"
/number=8
3291..3403
/gene="srp30"
/number=9
3404..3500
/gene="srp30"
/number=9
3501..3578
/gene="srp30"
/number=10
3579..4514
/gene="srp30"
/number=10
4178..4572
/gene="srp30"
/note="alternative exon"
/number=11
4178..4340
/gene="srp30"
/note="alternative exon 10a"
4515..4572
/gene="srp30"
/number=11
4573..4843
/gene="srp30"
/number=11
4844..5125
/number=12
BASE COUNT 1501 a 912 c 1038 g 1713 t
ORIGIN
Query Match 99.4%; Score 4021; DB 8; Length 5164;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 4043; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
QY 1 AGACAAAGATGCTTACTTCTTTAAACATGTTTCGAGTTTATGAAAATGATCACCAGCTTC 60
DB 1121 AGACAAAGATGCTTACTTCTTTAAACATGTTTCGAGTTTATGAAAATGATCACCAGCTTC 1180
QY 61 TAACTATGATCTCTTCTCCCTGCAAGCGAAGATGATGATTAACGGAATAT 120
DB 1181 TAACTATGATCTCTTCTCCCTGCAAGCGAAGATGATGATTAACGGAATAT 1240

3400	Db		GGAGGTCCTTTATTTTTTTTTTTTTTTTTCATAAACCTAGACATATAAGGGATTTTTATTTGT	3459
2341	Qy		AACCTATTTATGAAATTAACGTACTTCTAAATGCAATGCAGTGTGTCACCTGCTAGATCC	2400
3460	Db		AACCTATTTATGAAATTAACGTACTTCTAAATGCAATGCAGTGTGTCACCTGCTAGATCC	3519
2401	Qy		ATTTCCCGCGTTTACCGGCCCTTAGTCGTTCTCGCTCGCTATACAGCTCTGTCTCAAGG	2460
3520	Db		ATTTCCCGCGTTTACCGGCCCTTAGTCGTTCTCGCTCGCTATACAGCTCTGTCTCAAGG	3579
2461	Qy		TATGAGTGTAGATTTGTGTATCATATTATATATATGTAGTTACCCCTTCATGGATCACTTGT	2520
3580	Db		TATGAGTGTAGATTTGTGTATCATATTATATATATGTAGTTACCCCTTCATGGATCACTTGT	3639
2521	Qy		TC TTGCA TAGTAACCTCTTACTAGCTTTATTA CTTACAACTAAGCAACCTTTTGTGTCCT	2580
3640	Db		TC TTGCA TAGTAACCTCTTACTAGCTTTATTA CTTACAACTAAGCAACCTTTTGTGTCCT	3699
2581	Qy		CCGTACACAGTTGAAATTTGTTTGAGTCTTTTTTCCCTCATAGTGGACTAGTCTATTGTCA	2640
3700	Db		CCGTACACAGTTGAAATTTGTTTGAGTCTTTTTTCCCTCATAGTGGACTAGTCTATTGTCA	3759
2641	Qy		CTTGATTTTCTCTCTGTTGTGATGTTTTCTATGTCATGCAAACTCCAAATGAGGTAAAGG	2700
3760	Db		CTTGATTTTCTCTCTGTTGTGATGTTTTCTATGTCATGCAAACTCCAAATGAGGTAAAGG	3819
2701	Qy		TTACCTCCTGTGTTGGGATACAGAGTTCCTTTTTCATTTCTTACACGTGAATGTGTTG	2760
3820	Db		TTACCTCCTGTGTTGGGATACAGAGTTCCTTTTTCATTTCTTACACGTGAATGTGTTG	3879
2761	Qy		TTTTTTATGTTTTGAGTTCCTTGACAGAGATGCTCCCATCATATTAGTCCCTTTCCCTTTC	2820
3880	Db		TTTTTTATGTTTTGAGTTCCTTGACAGAGATGCTCCCATCATATTAGTCCCTTTCCCTTTC	3939
2821	Qy		TC TTG TGT CGTTCCTCTCTGGAAGTTTCCTTCGTATAAGCTTTTACTTCTTAACTTTTT	2880
3940	Db		TC TTG TGT CGTTCCTCTCTGGAAGTTTCCTTCGTATAAGCTTTTACTTCTTAACTTTTT	3999
2881	Qy		TCCAGCGACGGTGAAATTTATTATAGTATCAACCTCAATATCCGACCTATAATTTAAAGAAC	2940
4000	Db		TCCAGCGACGGTGAAATTTATTATAGTATCAACCTCAATATCCGACCTATAATTTAAAGAAC	4059
2941	Qy		ACTTAGCTAGATGTTCACTTTTGAAATTTATTTCTATCGGAAGGGCGAGATGATTTCT	3000
4060	Db		ACTTAGCTAGATGTTCACTTTTGAAATTTATTTCTATCGGAAGGGCGAGATGATTTCT	4119
3001	Qy		GAGGCATGTCCTCGATCATTTACCGTGATTAACAATTTGTA TTGCGTTGTTTGTGTTCTAGAT	3060
4120	Db		GAGGCATGTCCTCGATCATTTACCGTGATTAACAATTTGTA TTGCGTTGTTTGTGTTCTAGAT	4179
3061	Qy		CTGGCTCAGTCTACAGACTGGGGATGGATCTAGATGGGTCACTAGATGGGATTTCTGG	3120
4180	Db		CTGGCTCAGTCTACAGACTGGGGATGGATCTAGATGGGTCACTAGATGGGATTTCTGG	4239
3121	Qy		ACTGGATTTACAAGCTGGATTAGCATGAAC TGAACTTC TTGTTTACCGTCTCGTCTGGT	3180
4240	Db		ACTGGATTTACAAGCTGGATTAGCATGAAC TGAACTTC TTGTTTACCGTCTCGTCTGGT	4299
3181	Qy		CTGGTACTCCGCGGTATCAGCTGTAGGATCTGATCGCAAAGTTTTGGACTATGATTA CT	3240
4300	Db		CTGGTACTCCGCGGTATCAGCTGTAGGATCTGATCGCAAAGTTTTGGACTATGATTA CT	4359
3241	Qy		CTGANTCCTCAATATATATTATCTTTTGTGACAAATAGTGGATTTCTGTGTGAGTCTTTTCT	3300
4360	Db		CTGANTCCTCAATATATATTATCTTTTGTGACAAATAGTGGATTTCTGTGTGAGTCTTTTCT	4419
3301	Qy		AGGACAGCATTTAAGCTCCCGGACTAGATGGGAGATGGTCAGTAAATTTCTTTGTTATG	3360
4420	Db		AGGACAGCATTTAAGCTCCCGGACTAGATGGGAGATGGTCAGTAAATTTCTTTGTTATG	4479
3361	Qy		CCACACTTACATGGGGTTTTTTCGGTCTTGTGCGAGTCCCAATCAAGATCAAAATCAAGA	3420

4480	CCACATTACATGGGGTTTTTGGTCTTGGTCGAGCTGCCAATCAAGATCAAAATCAAGA	4539
3421	TCAGATCAAGATCGAAATCTCCAGTTTCCACCTGTGGTAAAGTCTAAAAGCTGAACCTTCT	3480
4540	TCAGATCAAGATCGAAATCTCCAGTTTCCACCTGTGGTAAAGTCTAAAAGCTGAACCTTCT	4599
3481	TTAAATTCAATCCATGTGTTTGGTTTAAATACCTGCTCACATTGGTGTCTTCTCAATCA	3539
4600	TTAAATTCAATCCATGGTGTGGTTTAAATACCTGCTCACATTGGTGTCTTCTCAATCA	4659
3540	ACACCAACTTAACGGAATCATGAGACAGACTATAAAATTTGAAGAGTCTGTGAACGACT	3599
4660	ACACCAACTTAACGGAATCATGAGACAGACTATAAAATTTGAAGAGTCTGTGAACGACT	4719
3600	AGTCTCACGAACCTCTGTGTGACACTAAAAATCGCCTCTCCAAAGTGTTCAGCAACATAA	3659
4720	AGTCTCACGAACCTCTGTGTGACACTAAAAATCGCCTCTCCAAAGTGTTCAGCAACATAA	4779
3660	TCTACCTCTGTCATGTGTTATCATTTCTTCTCTCTTAACGGTATTACATATTATGTTTT	3719
4780	TCTACCTCTGTCATGTGTTATCATTTCTTCTCTCTTAACGGTATTACATATTATGTTTT	4839
3720	GCAGGTGATATCTGGTTGAAAATGAAAACGTGGCCACTGGCTGTACCCGAATCGTCTCAAG	3779
4840	GCAGGTGATATCTGGTTGAAAATGAAAACGTGGCCACTGGCTGTACCCGAATCGTCTCAAG	4899
3780	CTTCTCAGGCTCCACTGCTAAATAGATTGATTCGGATTTGGGATTTATATCTGCTCTT	3839
4900	CTTCTCAGGCTCCACTGCTAAATAGATTGATTCGGATTTGGGATTTATATCTGCTCTT	4959
3840	CTTGTATGGACACCAATAATGTCCTTCTAGTTTTAGTTGTGAACCTGGAAATGGTCTGT	3899
4960	CTTGTATGGACACCAATAATGTCCTTCTAGTTTTAGTTGTGAACCTGGAAATGGTCTGT	5019
3900	TATGTGTCAATTAATAAGCCGGAACTCTGCTCGGCTGCATAATAAAGTTCAATCAGACA	3959
5020	TATGTGTCAATTAATAAGCCGGAACTCTGCTCGGCTGCATAATAAAGTTCAATCAGACA	5079
3960	TTGCTGTTGGTGTGGTGGAGTTTTCATACATATACATTTACATTAACAACCTACTCGGTGT	4019
5080	TTGCTGTTGGTGTGGTGGAGTTTTCATACATATACATTTACATTAACAACCTACTCGGTGT	5139
4020	CTTTTATGATATATCTTAAACTAAAC	4044
5140	CTTTTATGATATATCTTAAACTAAAC	5164

DECIIT. 3

T12M4

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS
POETRY

SOURCE
ORGANIZATION

TABLE 2

REFERENCE

AUTHOR.

AKPLMLTSSSTNSDISCLITRTALVGMVLYGGNMIDAHKVPFEMPERPVTHNV
MITGLNLGDFEKLALCELEKPNRTVVSWTIIIDGVARVDKPKREAILLPSRMVACDAI
KPEIITLAILPFAVNLGDMGSGVHAYVKGKGFPCDIFRVNSLIDAYAGCGCIQS
AFKFIPIBIPNGKSLVSWTMI SFAFHHGMEKAVSFMDKMERLGLKPNRVTMISVLN
ACSGHGLAESEFLEFFFTMVNVEKIPDVKHVGLVDMRLRGRLEAEAKI LEPIE
EKAVWMLLAGCSVYDDAEALERVTRKLMELERSHGGDVVLMSNFCTGTGRFLDAQR
FRQMDVRGKAPLPGHSQVT"
25594. .28968
/gene="T12M4.8"
join(25594. .25687,26293. .26400,26495. .26687,26785. .27043,
27217. .27321,27408. .27494,27577. .27634,28142. .28209,
28295. .28390,28457. .28531,28693. .28722,28807. .28968)
/gene="T12M4.8"
/note="Match to calreticulin (AtCRTL) mRNA gb|U27698 and
DNA gb|U66344. ESTs gb|T45719, gb|T22451, gb|H36323 and
gb|AA042519 come from this gene."
/codon_start=1
/evidence=not experimental
/protein_id="AAC24083.1"
/db_xref="GI:3249100"
/translations="MAKIPSLVSLILIGLVAISAAVIFERFDDGWMENRVKSEWK
KDNNTAGEMKHTAGNSGDANDKGIOTSEDRYFAISABFPFSNKDKTLVFOFVKH
EOKLDCGGVYKLLSGVDOKKFGDTPVSLNEGPDICGYSKKVHAILTYNEANHLI
KXVPESTDOLTHVYFILRPDAYSILIDNVEKQSGLYSDWDLPPKKIKDPSAKK
PEWDQOEYISDPEDKPKPDYDDIPKEIPDPSKPEDDDEDEGEWTAPTIPNPEYM
GEWKPKIKNPNTKKEWAPLIDNPDKDPELYVFPKLYGLELWVKSGLSLEDYV
LI CDDPYAKKLADETWGLKDNIOQNI TLQAEKAAFDEAEKKNVEESKDPAPSSDA
EDRPDEGGDDSDSESABETKSVDSSEBSEKDAHVCYHNI TWRFTHP"
30117. .30527
/gene="T12M4.9"

Query Match 86.4%; Score 3493.2; DB 8; Length 59261;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 3622; Conservative 0; Mismatches 8; Indels 14; Gaps 11;

Qy 1 AGACAAAGATGCTTACTTCTTAAACATGTTGAGGTTTATGAAAATGATCACCAGCTTC 60
Db 55624 AGACAAAGATGCTTACTTCTTAAACATGTTGAGGTTTATGAAAATGATCACCAGCTTC 55683

Qy 61 TAACTATGATGCTTCTTCCCTCCAGCGAACAGTGGAGATGATTGATAACGGAATAT 120
Db 55684 TAACTATGATGCTTCTTCCCTCCAGCGAACAGTGGAGATGATTGATAACGGAATAT 55743

Qy 121 CGGAACATCACTCAACAAACCAAAAATTTGGACATCATATCGCAACAAAATTCATAGGAA 180
Db 55744 CGGACATCACTCAACAAACCAAAAATTTGGACATCATATCGCAACAAAATTCATAGGAA 55803

Qy 181 AAATACCTGAATTCAAAAACAGAAAAACAAACCGAACAGAGCAGAACTCACGGACTGA 240
Db 55804 AAATACCTGAATTCAAAAACAGAAAAACAAACCGAACAGAGCAGAACTCACGGACTGA 55863

Qy 241 GAGACCGTGGACGGTGTACCGCGGAAAATGGTCTTGGACGGAGTTACTAATCGCGGA 300
Db 55864 GAGACCGTGGACGGTGTACCGCGGAAAATGGTCTTGGACGGAGTTACTAATCGCGGA 55923

Qy 301 ATTGAGATTTGAGAGGTGGTAGAGGAACCGAGAGATGTTCTCTCAAAAAATCCC 360
Db 55924 ATTGAGATTTGAGAGGTGGTAGAGGAACCGAGAGATGTTCTCTCAAAAAATCCC 55983

Qy 361 CAAGTGTTCGGATCTAGTGTCTCTTTTGTCCAAAAACGACAGTGTTTAGGAAACCTAGG 420
Db 55984 CAAGTGTTCGGATCTAGTGTCTCTTTTGTCCAAAAACGACAGTGTTTAGGAAACCTAGG 56043

Qy 421 AGAATGAATACCCGATGGTCCGAATCCGATTCGAAATGGTTGGAATTTGTAATACTAA 480
Db 56044 AGAATGAATACCCGATGGTCCGAATCCGATTCGAAATGGTTGGAATTTGTAATACTAA 56103

Qy 481 CATACAATATTCGGTTTGAATGATAGAAAAACACATCGATCCGGTTAGAACATAT 540
Db 56104 CATACAATATTCGGTTTGAATGATAGAAAAACACATTCGATCCGGTTAGAACATAT 56163

Qy 541 ATTAACAGGCCCAATTAACATATGGCGGATCTTGATCAACTGGGCTATTTCATCGTTCA 600
Db 57243 ATCTCTTTCTGAAATGAATGTTATTTCCTATATTTACCGGGTATTAGGATCACATCGCGCA 57302

56164 ATTAACAGGCCCAATTAACATATGGCGGATCTTGATCAACTGGGCTATTTCATCGTTGA 56223
Qy 601 TACATCGCGCGCACAGAGATTAAATCCAGTCCGTTTATTAATAAGGATACATAGTTTCCA 660
56224 TACATCGCGCGCACAGAGATTAAATCCAGTCCGTTTATTAATAAGGATACATAGTTTCCA 56283
Qy 661 AACGAAACCGTGGTGTCTCTCTTCCAGAACAAATCTAACAGCTTTCTCGAACATCTCTCT 720
56284 AACGAAACCGTGGTGTCTCTCTTCCAGAACAAATCTAACAGCTTTCTCGAACATCTCTCT 56343
Qy 721 TCTTCTTTCTCGAAATTAATTTTCCAGTAACTCAATTTCTCTCTCTCTCTCTCTCTCTCT 780
56344 TCTTCTTTCTCGAAATTAATTTTCCAGTAACTCAATTTCTCTCTCTCTCTCTCTCTCTCT 56403
Qy 781 AACTAAATTTCTGCTCTCAGGTATCAGATGATGACCGATGGAATCTGATCAGATCTACGTT 840
56404 AACTAAATTTCTGCTCTCAGGTATCAGATGATGACCGATGGAATCTGATCAGATCTACGTT 56463
Qy 841 GGGAAATTTCCCTGGAGATATTCGCAAGTGTGAGGTGAAGATCTCTCTCTCAAGGTTTGA 900
56464 GGGAAATTTCCCTGGAGATATTCGCAAGTGTGAGGTGAAGATCTCTCTCTCAAGGTTTGA 56523
Qy 901 AAAATTTCTCTCTCTCTCTCGATAAAATTTGAATTCATATGATATTTGTTTTCAGTA 960
56524 AAA-TTCTCTCTCTCTCTCGATAAAATTTGAATTCATATGATATTTGTTTTCAGTA 56582
Qy 961 TTTGCAATTTCTGCTCTGCTGAGACAAATTTAAATCGACTCTTATGATATTTGTTTTCAGTA 1020
56583 TTTGCAATTTCTGCTCTGCTGAGACAAATTTAAATCGACTCTTATGATATTTGTTTTCAGTA 56642
Qy 1021 TGGACCAATTTGGACATGATTTGAAGATTTCCACCGAGACCTCTCTGTTATGCTTTGT 1080
56643 TGGACCAATTTGGACATGATTTGAAGATTTCCACCGAGACCTCTCTGTTATGCTTTGT 56702
Qy 1081 CGAGTATATTTGATCAAGTACAAATTTGTTTTTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1140
56703 CGAGTATATTTGATCAAGTACAAATTTGTTTTTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 56762
Qy 1141 ATGACTAAGATAGTTGTTTATTTGTTGCGAGTTTGAAGATCTCTGATGACAGATGCA 1200
56763 ATGACTAAGATAGTTGTTTATTTGTTGCGAGTTTGAAGATCTCTGATGACAGATGCA 56822
Qy 1201 ATTTATGACGCTGATGTTTATGATTTTGGTGGTGTGACCTTCGCGTTAGTAAACGATG 1260
56823 ATTTATGACGCTGATGTTTATGATTTTGGTGGTGTGACCTTCGCGTTAGTAAACGATG 56882
Qy 1261 ATGAAAGCTAGCTTAAATTTCTGTAATTTCTGTTAAAGGTTTATCTTTGTTGATGTTT 1320
56883 ATGAAAGCTAGCTTAAATTTCTGTAATTTCTGTTAAAGGTTTATCTTTGTTGATGTTT 56942
Qy 1321 TTAGTTGAGATTCGACATGTTGTTGTTTCCACCATCAGTTGATAGGTACACAG 1380
56943 TTAGTTGAGATTCGACATGTTGTTGTTTCCACCATCAGTTGATAGGTACACAG 57002
Qy 1381 CAGCTACAGTCCGAGCGGTGACCTTCAAGACGCTCTGACTACCGCGTTTGTAGAGTCT 1440
57003 CAGCTACAGTCCGAGCGGTGACCTTCAAGACGCTCTGACTACCGCGTTTGTAGAGTCT 57062
Qy 1441 TCTCGATTTGTTTATTTGTTGTTTAAATTTTATATTTGAAAACTCAATTTTACTTAC 1500
57063 TCTCGATTTGTTTATTTGTTGTTTAAATTTTATATTTGAAAACTCAATTTTACTTAC 57122
Qy 1501 CTAACATGATGCTGTGTGACCGGATTCGCGCTTCTGCTTCGTGCGAGGACCTTAAG 1560
57123 CTAACATGATGCTGTGTGACCGGATTCGCGCTTCTGCTTCGTGCGAGGACCTTAAG 57182
Qy 1561 TAAGGACATATATAGTCTTTTCTCTGATGTTGTTCTCTATATCATGATTTTGGATT 1620
57183 TAAGGACATATATAGTCTTTTCTCTGATGTTGTTCTCTATATCATGATTTTGGATT 57242
Qy 1621 TATCTCTTTCTGAAATGATGTTTATTTGCTTATTTTACCGGGTATTAGGATCACATCGCGCA 1680
57243 TATCTCTTTCTGAAATGATGTTTATTTGCTTATTTTACCGGGTATTAGGATCACATCGCGCA 57302

[illegible]

Qy	2761	TTTTTTATGTTTTGAGTTCTTGACAGAGATGCTCCCATCATATTTAGTGCCTTTTCTCTTC	2820
Db	58382	TTTTTTATGTTTTGAGTTCTTGACAGAGATGCTCCCATCATATTTAGTGCCTTTTCTCTTC	58441
Qy	2821	TCCTTGCGTGGTCTCTCTGCGATGTTCCCTCTGATAAAGCTTTACTTCTTAACTTTT	2880
Db	58442	TCCTTGCGTGGTCTCTCTGCGATGTTCCCTCTGATAAAGCTTTACTTCTTAACTTTT	58501
Qy	2881	TCCAGCGACGGTGAAATTTATTACGTATCAA- COTCAATATCCGACCTATAATTTAAAGAA	2939
Db	58502	TCCAGCGACGGTGAAATTTATTACGTATCAATCTTCATATCCGACTATAATTTAAAGAA	58561
Qy	2940	CACCTAGCTAGATGTT- - CACTTTTGAAATTTATTTCTATCGGAAGGGGGCAGAT- GAT	2996
Db	58562	CACCTAGCTAGATGTTCTACTTTTGAAATTTATTTCTATCGGAAGGGGGAGATGGAT	58621
Qy	2997	TTCTGAGGCAATGCTCTCGATCATTTTACCGTGATTAACAATGTTATTTGCGTGTGTTTG- TTTC	3055
Db	58622	TTCTGAGGCAATGCTCTCGATCATTTTACCGTGATTAACAATGTTATTTGCGTGTGTTTGTTTC	58681
Qy	3056	TAGATCTGGCTCACTGCTACGAGCTGGGANTTGGATCTAGATGGGTCACTAGATGGGATT	3115
Db	58682	TAGATCTGGCTCACTGCTACGAGCTGGGANTTGGATCTAGATGGGTCACTAGATGGGATT	58741
Qy	3116	CTTGGACTGGANTTACAAAGCTGGATTAGCATGAAC- TGAACCTCTG- TTTTACGGTCTG	3173
Db	58742	CTTGGACTGGANTTACAAAGCTGGATTAGCATGAACCTTGAACTTTTACGGTCTG	58801
Qy	3174	GTCGTGCTGGTACTCCGGCGGTAACAGCTGTAGATCTGATCGCAAGTTTTTGGACTAT	3233
Db	58802	GTCGTGCTGGTACTCCGGCGGATCAGCTGTAGATCTGATCGCAAGTTTTTGGACTAT	58861
Qy	3234	GATTCTCTGATTCCTCAATATATTTATCTTTTGCATATCTGGATCTGCTGTGTAGTT	3293
Db	58862	GATTCTCTGATTCCTCAATATATTTATCTTTTGCATATCTGGATCTGCTGTGTAGTT	58921
Qy	3294	CTTTTCTAGGACAGCATTTAAGCTCCCGGACCTAGATGGGACATGGTCAGTAAATTTCTT	3353
Db	58922	CTTTTCTAGGACAGCATTTAAGCTCCCGGACCTAGATGGGACATGGTCAGTAAATTTCTT	58981
Qy	3354	TGTTATGCCACACTTACATGGGGTTTTCCGGTCTTGCTGCAGGTCCTCAATCAAGATCAAA	3413
Db	58982	TGTTATGCCACACTTACATGGGGTTTT- GGTCTTGCTGCAGGTCCTCAATCAAGATCAAA	59040
Qy	3414	ATCAAGATCAAGATCAAGATCGAATTCCTCCAGTTTCACCTGTGGTAAAGTCTAAAAGCTGA	3473
Db	59041	ATCAAGATCAAGATCAAGATCGAATTTCTCCAGTTTCACT- - GTAAGTCTAAAAGCTGA	59097
Qy	3474	ACCTTCTTAAATTCACATCCAT- GGTGTTGTTTAAATACCTGTCTCACTTTGGTGTCT	3532
Db	59098	ACCTTCTTAAATTCACATCCATGGTGTGTTTAAATACCTGTCTCACTTTGGTGTCT	59157
Qy	3533	TCAATCAACACCAACTTAACGAAATCATGAGACAGACTATAAAATTTGAAGAGTCTGTAG	3592
Db	59158	TCAATCAACACCAACTTAACGAAATCATGAGACAGACTATAAAATTTGAAGAGTCTGTAG	59217
Qy	3593	AACGACTAGTCTCAACAACTCTGTGTGCACTAAAATTCGGCT	3636
Db	59218	AACGACTAGTCTCAACAACTCTGTGTGCACTAAAATTCGGCT	59261

RESULT 4
AX509198
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Arabiopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1
 AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
 TITLE Stress-regulated genes of plants, transgenic plants containing same, and methods of use
 JOURNAL Patent: WO 021655-A 3993 28-FEB-2002;
 The Scripps Research Institute (US) ; Syngenta Participations AG (CH)
 FEATURES source
 1. .780
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 BASE COUNT 256 a 154 c 154 g 216 t
 ORIGIN
 Query Match 19.3% Score 780; DB 6; Length 780;
 Best Local Similarity 100.0%; Pred. No. 2.1e-161; Mismatches 0; Indels 0; Gaps 0;
 Matches 780; Conservative 0;
 28 GTTCAGGTTTATTGAAGAATGATCACCAGCTTCTTAATGATGTTATCTTCTCCCTCGAAG 87
 Db 1 GTTCAGGTTTATTGAAGAATGATCACCAGCTTCTTAATGATGTTATCTTCTCCCTCGAAG 60
 Qy 88 CGAACATGGAAGATGATTGATTAACCGGAATATCGGAACATCTCAACAAACCAAAAT 147
 Db 61 CGAACATGGAAGATGATTGATTAACCGGAATATCGGAACATCTCAACAAACCAAAAT 120
 Qy 148 TTGACATCATATCGCAACAAATCAATAGAAAATATCTGAATTTCCAAAACAGAAAAA 207
 Db 121 TTGACATCATATCGCAACAAATCAATAGAAAATATCTGAATTTCCAAAACAGAAAAA 180
 Qy 208 CCAACGGACAGAGAGGAACTCAGGACTGAGAGACCGTGGACGGTGTCAACGGCG 267
 Db 181 CCAACGGACAGAGAGGAACTCAGGACTGAGAGACCGTGGACGGTGTCAACGGCG 240
 Qy 268 AAAATGCTCTTGGACGGAGTTACTTAATCGGCGAATTGAGATTTGAGAGGTGGTAGAG 327
 Db 241 AAAATGCTCTTGGACGGAGTTACTTAATCGGCGAATTGAGATTTGAGAGGTGGTAGAG 300
 Qy 328 GAACCGAGAGATGTTCTCTCAAAAAATCCCAAGTGTTCGATCTAGTCTCTTT 387
 Db 301 GAACCGAGAGATGTTCTCTCAAAAAATCCCAAGTGTTCGATCTAGTCTCTTT 360
 Qy 388 TGTCAAAACGACAGTGTTTAGAAACCTAGAGATGAATGACCCGATGTCGGAATC 447
 Db 361 TGTCAAAACGACAGTGTTTAGAAACCTAGAGATGAATGACCCGATGTCGGAATC 420
 Qy 448 CGATTGCAAAATGGTTCGAATTGTAATACTAATTAACATACATATTCGGTTTGAATGATA 507
 Db 421 CGATTGCAAAATGGTTCGAATTGTAATACTAATTAACATACATATTCGGTTTGAATGATA 480
 Qy 508 GAAAAACACATTCGATCCGGTTAGAACATATATTAACAGGCCCAATTAACATATGGG 567
 Db 481 GAAAAACACATTCGATCCGGTTAGAACATATATTAACAGGCCCAATTAACATATGGG 540
 Qy 568 CCGATCTTGATCAACTGGGTATTCATCGTTGATACATGCGGCGGACAGGATTAATTC 627
 Db 541 CCGATCTTGATCAACTGGGTATTCATCGTTGATACATGCGGCGGACAGGATTAATTC 600
 Qy 628 CAGTCCGTTTATAAAGGATCTACTTTCCAAACGACGGTGTCTCTCTTCCAG 687
 Db 601 CAGTCCGTTTATAAAGGATCTACTTTCCAAACGACGGTGTCTCTCTTCCAG 660
 Qy 688 AACAACTAAGCGTTTCTCGAACATCTCTCTCTCTCTCTCTCTCTCGAAATTAATTTCCAG 747
 Db 661 AACAACTAAGCGTTTCTCGAACATCTCTCTCTCTCTCTCTCTCTCGAAATTAATTTCCAG 720
 Qy 748 TAATCAATTTCT 807
 Db 721 TAA'CAATTTCT 780

RESULT 5

F7G19 110514 bp DNA linear PLN 16-SEP-1997
 LOCUS Sequence of BAC F7G19 from Arabidopsis thaliana chromosome 1,
 DEFINITION complete sequence.
 ACCESSION AC000106
 VERSION AC000106.1 GI:2342673
 KEYWORDS HTG.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 REFERENCE 1 (bases 1 to 110514)
 Osborne, B.I., Vysotskaia, V.S., Toriumi, M., Yu, G., Oji, O.,
 Buehler, E., Conway, A.B., Conway, A.R., Dewar, K., Feng, J., Kim, C.,
 Kurtz, D., Li, Y., Shinn, P., Sun, H., Davis, R.W., Ecker, J.R.,
 Pederspiel, N.A. and Theologis, A.
 The sequence of BAC F7G19 from Arabidopsis thaliana chromosome 1
 Unpublished (1997)
 AUTHORS Theologis, A.
 TITLE Direct Submission
 JOURNAL Submitted (20-JAN-1997) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 REFERENCE 3 (bases 1 to 110514)
 Theologis, A.
 AUTHORS Theologis, A.
 TITLE Direct Submission
 JOURNAL Submitted (17-FEB-1997) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 REFERENCE 4 (bases 1 to 110514)
 Theologis, A.
 AUTHORS Theologis, A.
 TITLE Direct Submission
 JOURNAL Submitted (27-FEB-1997) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 REFERENCE 5 (bases 1 to 110514)
 Theologis, A.
 AUTHORS Theologis, A.
 TITLE Direct Submission
 JOURNAL Submitted (03-MAR-1997) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 REFERENCE 6 (bases 1 to 110514)
 Theologis, A.
 AUTHORS Theologis, A.
 TITLE Direct Submission
 JOURNAL Submitted (02-APR-1997) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 REFERENCE 7 (bases 1 to 110514)
 Theologis, A.
 AUTHORS Theologis, A.
 TITLE Direct Submission
 JOURNAL Submitted (16-APR-1997) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 REFERENCE 8 (bases 1 to 110514)
 Theologis, A.
 AUTHORS Theologis, A.
 TITLE Direct Submission
 JOURNAL Submitted (26-AUG-1997) Plant Gene Expression Center, 800 Buchanan
 St., Albany, CA 94710, USA
 REFERENCE 9 (bases 1 to 110514)
 Theologis, A.
 AUTHORS Theologis, A.
 TITLE Direct Submission
 JOURNAL Submitted (16-SEP-1997) Plant Gene Expression Center, 800 Buchanan
 St., Albany, CA 94710, USA
 COMMENT On Aug 26, 1997 this sequence version replaced gi:1841948.
 The sequence of BAC F7G19 from Arabidopsis thaliana chromosome 1.
 FEATURES source
 1. .110514
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="F7G19"
 971. .2850
 /gene="F7G19.1"
 join(971..1279,1540..1641,1698..1837,2134..2230,
 2556..2665,2740..2850)
 gene
 CDS

```

/genes="F7G19.1"
/notes="Similar to ATP-dependent Clp protease (gb|D90915).
EST gb|N65461 comes from this gene."
/codon_start=1
/evidence=not_experimental
/protein_id="AAB70396.1"
/db_xref="GI:2342674"
/translation="MASCIAQSMNSLLPRSSSPSPHPLSSSSGRRNLKTRYPAPRA
KASAKVIMPNRPKDPFLSLTAANSPEKLNRPNVADVPYLDIFSPOLMSSPA
QVRSVAPNRPRTPPDLPMSLDGRIVYGMPTDDDIIFCQLVPAYVELVAELM
YLOWLPKEFIYIINSTTRDDGTHTVCGAALGQALLSAGTGGKRFMPFHA
KGESBDFPIFILPVMQQLVITNRDILVELLSKTSVETVANYMRRPYMDAPKA
KEFVSIDRVSLGTGVI"
complement(3700..3771)
/product="tRNA-Glu"
/notes="codon recognized: GAA"
/genes="F7G19.2"
join(3942..4046,4391..4507,4747..4854,4973..5164,
5263..5454,5551..5669,5747..5834,5916..6035,6216..6356,
6436..6513)
/genes="F7G19.2"
/notes="Similar to probable Mg-dependent ATPase
(pir|556671). ESTs gb|T46782.gb|AA04798 come from this
gene."
/codon_start=1
/evidence=not_experimental
/protein_id="AAB70397.1"
/db_xref="GI:2342675"
/translation="MAEDRSFEGDQLASMTTDDIGRASLLANEIRILKEESORTNLD
LESVKIKENQSKIKNLQPLVNLGNIVLELMSPEDDAEDGANIDLDORKGCV
VLKTSRQIFLPLVGLPDPDLKPGELGVNKSXYLLDITLPSYDSRVKAMEVDEK
PTDYNDIGLEKIQIELVEAIVLPMTHEKQDGIIRPGKGVLYGGPTGKILMAR
ACAAQTNAPFLQAGLVQMGFSDGAKLVDAFLAKESPCIIFIDEIDAIKTRF
DSVSGDREVRQTMELLQDLFGSDRIKVIATNRADILDPALMSGRDLRIEF
PHTEARGLIQLHRSKKNVADNVFELARSTDDFNGAQLKAVCVGAEMLALRDA
TEVNHEDNEGIIQVQAKKASLYYA"
complement(6693..9149)
/genes="F7G19.3"
complement(join(6693..6802,6892..7391,7471..7544,
7619..7721,7805..7920,8032..8127,8253..8639,8799..9149))
/genes="F7G19.3"
/notes="Strong similarity to Oryza NADPH oxidase
(gp|X93301)."
/codon_start=1
/evidence=not_experimental
/protein_id="AAB70398.1"
/db_xref="GI:2342676"
/translation="MUSQLIPKDRNPVKRFAMNISYFFLENKRIWLVLTWISICI
TLFTWFLQYKRTVEVNGYCVTVAKSAETLKFMALILLPVCNRTITWLTKSL
IGSVFPDDNINFHKVAFGIAGVIGLHAIHLACDFPRLHAKVVEPEPMKFFGDE
RPNYGMFKGTGDMGTVMVLMVAVYLAQSWFRRNRANLPKSLKLTGTFNAFWS
HLFVTVVLLIVHGVFVYLSKEWYHKTVMYLAVPVLLYAFERLIRAPRGAKAVY
LKAVYVGNVLSLYMSKPKGPKYTSQYIYINCSDVSPLOWHPFSITASGDYLSVH
IRTLGWTLSQKLSKVCQLPSTQSGLFIADIGQANNITFPRLLDIGPYGAPAQ
YRNDVLLVGLIGATPLISIRIDVNLKNQNSIERGNTQHNRYATKRAYFYW
TRQGLSEFVSWNEVAYDSBGMIELHNYCTSYVEEGDARSALITLMQSLHAKSG
IDIVSTVRVTRFARPNWRSVFHVAHVNVQVGVYCGNCTIIGELKRLAQDFSRK
TITTKFEHKNF"
complement(9310..10430)
/genes="F7G19.4"
complement(join(9310..9492,9582..9630,9745..9904,
10061..10430))
/genes="F7G19.4"
/codon_start=1
/evidence=not_experimental
/product="F7G19.4"
/protein_id="AAB70399.1"
/db_xref="GI:2342677"
/translation="MESSSEGTNKSIRCKATGSDNPDEYVEITLEVDERINTMKA
KATRLVSLGRKLVKWSUSFASRRDRSKSGFAMFALRGLRPIAKNDAVAGWDEVA
MRFDKLVAGPKSKFGICIGVSESFEVNFELVIRGRTGTSSTIKTLEFFWE
QITGNSFDRQLQIFDMVDKNDLGRITGDEVKEIIALSASANKLSKINENVDEYALI

```

```

MEELDRNLGYIEVRCLNSNGYTCLTSLKLTWLNPN"
/genes="F7G19.5"
join(11553..11812,11908..12119,12289..13011,13083..13344,
13419..13929)
/genes="F7G19.5"
/notes="Similar to Arabidopsis luminal binding protein
(gb|D89342)."
/codon_start=1
/evidence=not_experimental
/protein_id="AAB70400.1"
/db_xref="GI:2342678"
/translation="WSSLAIEGERQKLTGIVIGLTTYSVGVYHVKHVEIIANDQG
NRITPSVAFTDTERLGEAKNOAKNPRTIPDKRLIGRKFPDVPDVIDIKFLPY
KYVNDKQKPIQVVKGEELKFSPEELSAMLTWKETAFAFLGKKIKDAVITVPAYF
NDAQRQATKQAGATAGLVNRIINEPTGAALAYGLDKKGGESNLVLDLGGGTFDVS
LTIINDGVFVLSGTDETHLGGEDFHRVMDYFIKLVKKKYNDIKFTHKALGKLRREC
ELAKRSLNQHVRETESLFDVDFSEPLTRAFELNLDLKKTHKALGKLRDAG
LKSDIDBIVLVGGSTRIPKVOOMLKDFPFCKGSKGTNPDEAVYAKKASOVGVLSGE
GGETQKNLTLLKTDILLDVAPLSLGIEVTVGGVMTNLIIPRTVPIPKKSVFTTYQD
QQTVTNVGEGRSMYKDNRELKQFDLTGLIPAPRNQYFRGVQIEVTFEVDANGIL
QVKAEDVAKTSQSITITNDKRLTESEIEIREAEFEAEEDKIMEKIDARKNLET
YVYNKSTVADKEXLAKKISDEDEKNEGVLKEALEWLENVNAEKEDYDEKLEKVEL
VCDPVIKSVYKTEGENEDTWMETMSYRIV"
complement(14446..15420)
/genes="F7G19.6"
complement(14446..15420)
/genes="F7G19.6"
/notes="Similar to Glycine SRC2 (gb|AB000130). ESTs
gb|H76969.gb|T21700.gb|ATTS5089 come from this gene."
/codon_start=1
/evidence=not_experimental
/protein_id="AAB70401.1"
/db_xref="GI:1922937"
/translation="MECRSLDLTIIASBDLKDVLQIKQDLVAVVSINGDARTKQTK
VDKCGTKPKWQHOMKLTVDAAARDNRLTLVPEIVADRPAGDKPKVGEVSVPKELL
DQNKDSEKVTYAVRLPNKAKGSLKSEKFEKTYTGGSSGPHAPVPSMDKHTWD
QPVTAYPPGAPSAYPAPFAGPSGYPFGQDDHDGVTGYGPOQAGYPPAGTGYPFP
GAYPQGGYPPYPPQGGYPPYPPQGGYPPYPPQGGYPPYPPQGGYPPYPPQGGYPP
HGKAGAGMGLGLGILGAGLLGGLVGEAVSDIADMGDMGDMGDMGDF"
17933..21752
/genes="F7G19.7"
join(17933..18538,18639..18781,18879..19029,19082..19181,
19238..19367,19470..20115,20194..20523,20583..20618,
20661..21032,21125..21194,21284..21369,21470..21577,
21678..21752)
/genes="F7G19.7"
/notes="Similar to Vicia sativa ENBP1 (gb|X95995)."
/codon_start=1
/evidence=not_experimental
/protein_id="AAB70402.1"
/db_xref="GI:2342679"
/translation="MQVNFDETCDSVIRMNQETRSANGIGNGSGESIPGIPDDLRC
KRSQKQWRCTMSMADKTYCEKHVIOAKRAANSAPRANQKQKAKRSESSIGEDTYS
GKMDDFELPVTSIDHYNNGLASKSNRLEKHKHNSRISYFETPMWRSFSPKVAVD

```

Query Match 14.1%; Score 571; DB 8; Length 110514;
Best Local Similarity 99.2%; Pred. No. 2.1e-115;
Matches 606; Conservative 0; Mismatches 0; Indels 5; Gaps 3;

QY	3435	GAATTCCTCCAGTTTCACCTGGTAACTCTAAAGCTGAAACCTTCTTTAATTCAATCC	3494
Db	1	GAATTCCTCCAGTTTCACCTCT--GTAAGCTAAAGAGCTGAACCTTCTTTAATTCAATCC	57
QY	3495	AT-GTGTGTTTAAATACCTGCTCAGTTGGTGTCTTCAATCAACACCACTTAACG	3553
Db	58	ATGTGTGTTTAAATACCTGCTCAGTTGGTGTCTTCAATCAACACCACTTAACG	117
QY	3554	AAATCATGAGACACTATATAATTTTGAAGAGCTCTGTAGAACGCTAGGTCTTCCACCAACC	3613
Db	118	AAATCATGAGACACTATATAATTTTGAAGAGCTCT-TAGAACGCTAGGTCTTCCACCAACC	176
QY	3614	TCTGTGTGCATAAATAATCGCTCTCCCAAGTGTTCAGAACATAATCTACCTCTGTCAT	3673

```

Db      177 TCTGTGTGCACTAAAAATCGCTCTCCAAAGTGTTCAGCAACATAATCTACCTCTGTCAAT 236
Qy      3674 GTGTATCATTTCTCTCTCTTAACGGTATTACATATTATGTTTGCAGGTGATATCTG 3733
Db      237 GTGTATCATTTCTCTCTCTTAACGGTATTACATATTATGTTTGCAGGTGATATCTG 296
Qy      3734 GTTGAATAAGAAACTGGCCACTGGCTGATACCCGAATCGTCTCAAGCTTCTCAGGCTCCA 3793
Db      297 GTTGAATAAGAAACTGGCCACTGGCTGATACCCGAATCGTCTCAAGCTTCTCAGGCTCCA 356
Qy      3794 CTGCTAATAGAAATTTGATTCGGATTGGGATTAATACCTGCTCTCTTCTTATGGGACGA 3853
Db      357 CTGCTAATAGAAATTTGATTCGGATTGGGATTAATACCTGCTCTCTTCTTATGGGACGA 416
Qy      3854 CCAATATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3913
Db      417 CCAATATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 476
Qy      3914 AAACCGGAAACTCTGCTCGCTGCATATAAAGTTTCATCAGACATTTGTTGGGTGTG 3973
Db      477 AAACCGGAAACTCTGCTCGCTGCATATAAAGTTTCATCAGACATTTGTTGGGTGTG 536
Qy      3974 GTGAGGTTTTCCATACATATACATTTACATTTACATTTACATTTACATTTACATTTATC 4033
Db      537 GTGAGGTTTTCCATACATATACATTTACATTTACATTTACATTTACATTTACATTTATC 596
Qy      4034 TTAACCTAAAC 4044
Db      597 TTAACCTAAAC 607

RESULT 6
LOCUS   AY050912
DEFINITION Arabidopsis thaliana putative SP2/ASF splicing modulator Srp30
ACCESSION AY050912
VERSION   AY050912.1
KEYWORDS FLI CDNA.
SOURCE   Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 933)
Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,
Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K.,
Davis,R.W., Ecker,J.R. and Theologis,A.
Arabidopsis Full Length cDNA Clones
Unpublished
2 (bases 1 to 933)
Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Banno,F., Dale,J.M.,
Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L.,
Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C.,
Kosena,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
Southwick,A., Tracy,S.E., Shinozaki,K., Davis,R.W., Ecker,J.R. and
Theologis,A.
Direct Submission
Submitted (24-JUL-2001) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA'); Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

```

The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Banno,F., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Kosena,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tracy,S.E., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGSC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGSC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES

```

source
1..933
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAFL08-12-105 (R11038)"
/note="This clone is in a modified pBluescript vector (pLC-1) as a BamHI/XhoI insert."
ecotype: Columbia
<1..933
/gene="At1g09140"
<1..626
/gene="At1g09140"
/codon_start=3
/evidence=experimental
/product="putative SP2/ASF splicing modulator Srp30 protein"
/protein_id="AAK93589.2"
/db_xref="GI:21954085"
/translation="IYGRDGYDFGCRLEIARHGRRFSPVDYSSYSASRAPSR
RSDHVLVTGLPPASWQDLKDMRKAGDVCSEVFPDRKMGVWVDYNDVMKYAI
RKLDATERNAFSSAYINRYEYSRVSRRSPDDSKSYSRSRGSPGCSYSSKRSVS
FARSISPRSFLSRSLYSVSRSQSKSKRTKRSNSPVSPIVG"
misc_difference 1
/gene="At1g09140"
/note="compared to genomic sequence"
/misc_difference 148
/gene="At1g09140"
/replace="c"
/misc_difference 579
/gene="At1g09140"
/replace="g"
/note="compared to genomic sequence resulting in an amino acid sequence difference"
/replace="t"
627..933
/gene="At1g09140"
BASE COUNT 248 a 187 c 223 g 275 t
ORIGIN
Query Match 7.7%; Score 311.4; DB 8; Length 933;
Best Local Similarity 96.7%; Pred. No. 5.5e-58;
Matches 318; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
3715 GTTTTCAGGTGATATCTGTTTGAATAAATGAAACCTGGCCACTGGCTGTACCCGAATCGTC 3774
603 GTTTCACCTGTGATATCTGTTTGAATAAATGAAACCTGGCCACTGGCTGTACCCGAATCGTC 662
3775 TCAAGTTCTCAGGCTCCACTGCTAATAGATTTGATTCGATTTGGGATTTATATCTG 3834
663 TCAAGTTCTCAGGCTCCACTGCTAATAGATTTGATTCGATTTGGGATTTATATCTG 722
3835 GTCTTCTGATGGACGACCAATATGCTTCTAGTTTTAGTTTCTGAACCTGGAAATGG 3894
723 GTCTTCTGATGGGACGACCAATATGCTTCTAGTTTTAGTTTCTGAACCTGGAAATGG 782

```

```
QY 3895 TCTGTTATTGTGTCATTAATAAGCCGAAACTCTGTCCTGGGTGCATATAAAGTTTCATC 3954
Db 783 TCTGTTATTGTGTCATTAATAAGCCGAAACTCTGTCCTGGGTGCATATAAAGTTTCATC 842
QY 3955 AGACATTGTTGGGTGTCGTTTTCATACATATACATTACATTACAACTACT 4014
Db 843 AGACATTGTTGGGTGTCGTTTTCATACATATACATTACATTACAACTACT 902
QY 4015 GGTGCTTTTATGATTATCTTAACTAAA 4043
Db 903 GGTGCTTTTATGATTATCTTAACTAAA 931

RESULT 7
AP006409/c 109465 bp DNA linear PLN 11-JUN-2003
LOCUS Lotus japonicus genomic DNA, chromosome 5, clone:ljT25E08, TM0299,
DEFINITION complete sequence.
ACCESSION AP006409
VERSION AP006409.1 GI:31581040
KEYWORDS HTG.
SOURCE Lotus japonicus
ORGANISM Lotus japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteeae;
Lotus.
REFERENCE 1
AUTHORS Asamizu, E., Kato, T., Sato, S., Nakamura, Y., Kaneko, T. and Tabata, S.
TITLE Structural Analysis of a Lotus japonicus Genome. IV. Sequence
Features and Mapping of seventy-three TAC clones which cover the
7.5 Mb Regions of the Genome
JOURNAL DNA Res. (2003) In press
REFERENCE 2 (bases 1 to 109465)
AUTHORS Sato, S.
TITLE Direct Submission
JOURNAL Submitted (07-MAY-2003) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba 292-0818, Japan (E-mail: ssato@kazusa.or.jp,
URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935 (ex. 2337),
Fax: 81-438-52-3934)
FEATURES
    source
        Location/Qualifiers
            1..109465
                /organism="Lotus japonicus"
                /mol_type="genomic DNA"
                /db_xref="taxon:34305"
                /chromosome="5"
                /clone="ljT25E08"
                /clone_lib="ljT library"
                /notes="TAC clone: TM0299"
BASE COUNT 35258 a 19989 c 19616 g 34602 t
ORIGIN
Query Match 3.4%; Score 139; DB 8; Length 109465;
Best Local Similarity 54.6%; Pred. No. 4.7e-20;
Matches 481; Conservative 0; Mismatches 295; Indels 105; Gaps 6;
QY 991 AATCGACTCTATGATATATTTGTTTCAGATGAGCAATTTGGACATTTGAAGAT 1050
Db 92854 AATTGATTTCTTTTGTCTAATATCAGTTTGGTCTCTATTGTTGACATTTGAAGAT 92795
QY 1051 TCCACCGAGACCTCTGGTTATGCTTTTTCGAGGATATTT-----G 1092
Db 92794 TCTCTCAAGACCAACAGGATATGCTTTTGTTCAGGATGTTCTTTGAACTATTTTGGCGAG 92735
QY 1093 ATCAAGTACAAATTTGTTTTTTTCTTCTTCTTCTTAATAGTATAGGCTAATGACTAAGATA 1152
Db 92734 GACATGAGAGATTTATCATTTATCATTCACGTGAGTTTAGTCGCTTCTTCTGCTTTATA 92675
QY 1153 GTTTGTTATTGGTG-----CGAGTTTGAAGATCTCTGTTGATGAGCAGATGCAATTTATGG 1208
Db 92674 TTTATCAAAATGTTATATTGTCAGTTTGAGGATGTCGTGATGCTGAAGATGCAATCTACTA 92615
```

```
QY 1209 ACGTGATGGTTATGATTTTGTGACCTTCGACTTCGGGTTAGTAAAGCATGATAAGC 1268
Db 92614 TCGAGATGGTTATGATTTTGTGATGGTTTCGATACGAGTTAGTACACTTACCCCTCCITAC 92555
QY 1269 TAGCTTAATTTTCGTAATTTCTTTGTA-----AAGGTGTTATCTT 1308
Db 92554 CTTTTCGCTTTTGGTCTCTATGGGACTCACTATCTAAATTTGCTATAAAACTTTTGGTCA 92495
QY 1309 TGTGCTGATGTTTTAGGTTGAGATTGCACATGGTGGTCGTAGATTTTTCACCATCAGTTGA 1368
Db 92494 TTGCTTATTACTGTAGGTTGAACCTTGACATGGTGGACGGGGTTCTTTCATCATCAGTAGA 92435
QY 1369 TAGGTACAG-----CAGCAGCTACAGTGGGAGCGTGACCTTCAAGACGCTCTGACTA 1422
Db 92434 CCGTTACAGTAGTTATAGTGTGCGCAGCAGTAGCCGTGGAGCTTCCAGAGGGTCTGACTA 92375
QY 1423 CCGCGGTTTGTAGAGTCTTCTCGCATTTGTTATTTGGTGTGTGTAATAAATTTTATATTG 1482
Db 92374 TCGTGGTATGTTGTTTCCAAATATCAGTTTTTTGTAGCTTGATATCATATATAGTTA 92315
QY 1483 AAAACTCATTTTACTACCTAAACA----- 1507
Db 92314 TTTATCTTATTTCTTTTGTCTTATTTTGTCTTTGTATCATCAAGAGAAAATTTTCAC 92255
QY 1508 -----TGTAGTCTTGTGACCGGATACCGCTTCTGCTTCTGTTGG 1547
Db 92254 TGACCCATTATTGTTCTGTTGTAGTTCTGGTGAATGATCTCTTCTGCTTCAAGG 92195
QY 1548 CAGGACCTTAAGGTAAGGACACTATATAGTCTTTT--TTCTCTGAATGTGGTTCTCTAT 1605
Db 92194 CAGACCTGAAAGTATGATGATGCTTCTTCTTTTACTACTGAAAGTATGATGTTAAA 92135
QY 1606 ATCATGTTTGGATTTATCTCTTTTCTGAAATGATGTTATTGCTATTACGGGTGATTA 1665
Db 92134 TTATTGTTGTTTAAATAATAGTGAACAATATTGGCTGATTTAGTATTTAAATGACTA 92075
QY 1666 GGATCATCATGGCAAGCTGAGATGCTGCTCTCTGAAAGTTTTCCTGACCGTAAGG 1725
Db 92074 GGATCATCATGGTAAGCTGATGCTGCTCTCTCAAGTTTTCGCTGAGCGTGGAGG 92015
QY 1726 TGACTTGACATTGCGATAGTTTGGATAGCTTTTGTATTGAT 1766
Db 92014 TAAGTGATTGAGATATAATATTAGCAGAAATTTTATTGAT 91974

RESULT 8
AC144541/c 133656 bp DNA linear HTG 12-JUN-2003
LOCUS Medicago truncatula clone mth2-28a1, WORKING DRAFT SEQUENCE, 7
DEFINITION unordered pieces.
ACCESSION AC144541
VERSION AC144541.5 GI:31621362
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 133656)
AUTHORS Shaull, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
TITLE Medicago truncatula BAC Clone mth2-28a1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 133656)
AUTHORS Shaull, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (29-APR-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 820 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 3 (bases 1 to 133656)
```


	Matches	123;	Conservative	0;	Mismatches	10;	Indels	0;	Gaps	0;
Qy	2152	TCTCTGAATTA	TACTGTTTCAGT	CAGTGAGGAAATATGATCGAGAGGTGTCAGTCGAAGCCACG	2211					
Db	524	TCTCTAGTGC	TATTATACCGSGTGAGGAAATATGATCGAGAGGTGTCAGTCGAAGCCACG	583						
Qy	2212	ATGATTTCTAAAAGC	TATAGAAGCAGGAGTCGGAGCCGTGGTCCAAGCTGTAGCTATAGTA	2271						
Db	584	ATGATTTCTAAAAGC	TATAGAAGCAGGAGTCGGAGCCGTGGTCCAAGCTGTAGCTATAGTA	643						
Qy	2272	GCAAGACGACGAG	2284							
Db	644	GCAAGACGACGAG	656							

RESULT 10				
AY128356/c				
LOCUS	AY128356	939 bp	mRNA	linear
DEFINITION	Arabidopsis thaliana unknown protein (Atlg09150) mRNA, complete cds.			

ACCESSION	AY128356
VERSION	AY128356.1 GI:22135953
KEYWORDS	FLI CDNA.
SOURCE	<i>Arabidopsis thaliana</i> (thale cress)
ORGANISM	<i>Arabidopsis thaliana</i>
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	1 (bases 1 to 939)
AUTHORS	Tripp, M., Southwick, A., Karlin-Neumann, G., Nguyen, M., Miranda, M., Palm, C.J., Bowser, L., Jones, T., Banh, J., Carninci, P., Chen, H., Cheuk, R., Chung, M.K., Hayashizaki, Y., Tshida, J., Kamiya, A., Kawai, J., Kim, C., Lin, J., Liu, S.X., Narusaka, M., Pham, P.K., Sakano, H., Sakurai, T., Satou, M., Seki, M., Shinn, P., Yamada, K., Shinzaki, K., Ecker, J., Theologis, A. and Davis, R.W.

TITLE Direct Submission
JOURNAL Submitted (01-JUL-2002) DNA Sequencing and Technology Center,
Stanford University, 955 California Avenue, Palo Alto, CA 94304,
USA
COMMENT e-mail for correspondence: arab@sequence.stanford.edu

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'AIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y., and Shinozaki, K.

The Salk, Stanford, PGEc (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAPL cDNAs: Tripp, M., Nguyen, M., Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C. J., Bomser, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M. K., Kim, C., Lin, J., Liu, S. X., Pham, P. K., Sakano, H., Shinn, P., Yamada, K., Ecker, J., Theologis, A. and Davis, R. W.

Tripp, M. (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R. W. (SSP/Stanford) contributed equally to this work as PIs.

```

FEATURES             Location/Qualifiers
     1..939
         organism="Arabidopsis thaliana"
         mol_type="mRNA"
         db_xref="taxon:3702"
         chromosome="1"
         clone="RAFL04-17-K04 (R15763)"
         notes="This clone is in pBluescript
ecotype: Columbia"
     1..939
         gene="At1g09150"
         166..711
             gene="At1g09150"
             codon_start=1
             product="unknown protein"
             protein_id="AA091559.1"

```

```

/db_xref="GI:22135954"
/translation="MPKFCLEBIEISQNVKASVQRIRQSTQDEYPGLESVWEDLLP
KKPIIVKPCNHLTLVVNNYLPFCIRDGPYMLRLHLYPNIMKRFQVDRGAIK
FVLSGNNMCPGLTSPGVDQVEVERAEFVAIEGQKHALAIGFTPMKAKDIKSIN
GIGVNNHYINDGLAKMERLD"
276 a 204 c 194 g 265 t
BASE COUNT
ORIGIN

```

	Query Match	2.8%;	Score 113.8;	DB 8;	Length 939;
	Best Local Similarity	98.3%;	Pred. No. 2.3e-14;		
	Matches 115;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	229	CTCAGGACTGAGAGAGACCGTGGACGGTGTCTACGGCGGAAAATGCTCTTGGACGGAGTT	288		
DB	117	CTTCCGACTTGAGAGAGACCGTGGACGGTGTCTACGGCGGAAAATGCTCTTGGACGGAGTT	58		
QY	289	ACTAATCGGCGAATTGAGATTTGAGAGTGGTAGTAGAGAACCGAGAGATGTTTC	345		
DB	57	ACTAATCGGCGAATTGAGATTTGAGAGTGGTAGTAGAGAACCGAGAGATGTTTC	1		

[illegible]

REFERENCE

1. Cossius, eucorpus 17; Brassicaceae, Brassicaceae, Arabidopsis.

AUTHORS

Harper,J.F., Krops,J., Wang,X. and Zhu,T.

TITLE

Stress-regulated genes of plants, transgenic plants containing same, and methods of use

JOURNAL

Patent: WO 0216655-A 1199 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)

FEATURES	Location/Qualifiers
source	1. .762

BASE COUNT	189 a	161 c	197 g	215 t
ORIGIN				

	Query Match	2.7%	Score 108.8	DB 6	Length 762
	Best Local Similarity	90.6%	Pred. No. 2.9e-13		
	Matches 116	Conservative 0	Mismatches 12	Indels 0	Gaps 0
Qy	1300	TGTAICTTTGTCTGATGTTTTAGTTTGAGATTGCACATGGTGGTCTGATATTTTCACC	1359		
Db	204	TGATTTTGATGGGTGTGCACTTCGGGTGAGATTGCACATGGTGGTCTGATATTTTCACC	263		
Qy	1360	ATCAGTTGTATAGTACAGCAGCAGCTACAGTGGCAGCGCTGCACCTTTCAGACGCTCTGA	1419		
Db	264	ATCAGTTGTATAGTACAGCAGCAGCTACAGTGGCAGCGCTGCACCTTTCAGACGCTCTGA	323		
Qy	1420	CTACCGCG	1427		
Db	324	CTACCGCG	331		

RESULT 12	AC126791/c	AC126791	120374 bp	DNA	linear	PLN 24-APR-2003
LOCUS		Medicago				
DEFINITION		truncatula		clone mth2-7k2,		complete sequence.
ACCESSION		AC126791				
VERSION		AC126791.16	GI:29650264			
KEYWORDS		HTG.				
SOURCE		Medicago truncatula		(barrel medic)		


```

ORGANISM      Medicago truncatula
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               rosids; eustersids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
               Medicago.
REFERENCE      1 (bases 1 to 120374)
AUTHORS       Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
               Cook,D., Kim,D. and Roe,B.A.
TITLE         Medicago truncatula BAC Clone mth2-7k2
JOURNAL       Unpublished
REFERENCE      2 (bases 1 to 120374)
AUTHORS       Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
               Cook,D., Kim,D. and Roe,B.A.
TITLE         Direct Submission
JOURNAL       Submitted (03-JUN-2002) Department Of Chemistry And Biochemistry,
               The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
               OK 73019, USA
REFERENCE      3 (bases 1 to 120374)
AUTHORS       Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
               Cook,D., Kim,D. and Roe,B.A.
TITLE         Direct Submission
JOURNAL       Submitted (19-MAR-2003) Department Of Chemistry And Biochemistry,
               The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
               OK 73019, USA
REFERENCE      4 (bases 1 to 120374)
AUTHORS       Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
               Cook,D., Kim,D. and Roe,B.A.
TITLE         Direct Submission
JOURNAL       Submitted (09-APR-2003) Department Of Chemistry And Biochemistry,
               The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
               OK 73019, USA
REFERENCE      5 (bases 1 to 120374)
AUTHORS       Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
               Cook,D., Kim,D. and Roe,B.A.
TITLE         Direct Submission
JOURNAL       Submitted (10-APR-2003) Department Of Chemistry And Biochemistry,
               The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
               OK 73019, USA
REFERENCE      6 (bases 1 to 120374)
AUTHORS       Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
               Cook,D., Kim,D. and Roe,B.A.
TITLE         Direct Submission
JOURNAL       Submitted (24-APR-2003) Department Of Chemistry And Biochemistry,
               The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
               OK 73019, USA
COMMENT       On Apr 9, 2003 this sequence version replaced gi:29124158.
               ----- Genome Center
               Center: Department Of Chemistry And Biochemistry
               The University Of Oklahoma
               Center code:UOKNOR
               -----
FEATURES             Location/Qualifiers
     source           1..120374
                     /organism="Medicago truncatula"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3880"
                     /clone="mth2-7k2"
                     /clone_lib="Medicago truncatula BAC library H1"
                     /note="This is one of two clones in the same well from
                     mth2-7k2"
BASE COUNT      40119 a 19397 c 19969 g 40869 t
ORIGIN
Query Match      2.7%; Score 107.6; DB 8; Length 120374;
Best Local Similarity 57.1%; Pred. No. 3.9e-13;
Matches 298; Conservative 0; Mismatches 184; Indels 40; Gaps 4;
QY      1014 TTCAGTATGGACCAATTGGACATTGATTGAAGATTCCACCGAGACTCTCTGTTATG 1073
DB      70902 TCCAGTATGGTCTTACATTGATTGAGATCCCTCCGAACACCACGAGTTATG 70843
QY      1074 CTTTGTGCGAGTAT-----ATTGATCAAGTACAAATTGTTTCTTCTTC 1122
               ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB      70842 CTTTGTGAGGTATGTCTACATTTTCATTTTAAAGATGCGAAGGGGTAGTTATCACAC 70783
QY      1123 TTGTAATAGTA-----TAGGCTAATGACTAAGATAGTTTGTATTGTTGGCAGCTTTGA 1175
DB      70782 TTTTGTAGTTCCCTTGTGTTCTGTCTGTCTAATGATCTTAACATATATTGCACTTTGA 70723
QY      1176 AGATCCTCTGTATGACGACGATGCAATTTATGGACGTGATGTTATGATTTTGTATGGGTG 1235
DB      70722 GATGCTGCTGATGCTCAAGATGCGATTATTATAGAGATGCTTATGATTTTGTATGTTGA 70663
QY      1236 TCGATCTCGGTTAGTAAACGATGATGAAGCTAGCTTAATTTCTCTTAATTTCTTCTGA 1295
DB      70662 TCGATTACTGGTTAGAATATTTTCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 70603
QY      1296 AAGTGTATCTTTCTGTGATGTT-----TTTAGGTTGAGATTTCCACAT 1339
DB      70602 TCGAATTTTGTCAATGACCTTTTGGTCACTATTTGTTACTGTAGTTGAACTTCCACAT 70543
QY      1340 GGTGTCGTAGATTTTTCACCATCAGTTGATAGTTACAG-----CAGCAGCTACAGTGGG 1393
DB      70542 GGTGACGGGGAAGTTTCATCATCAGTACAGCATAGGATAGCGGTAGAGTGGT 70483
QY      1394 AGCCGTGACCTTCAAGACGCTCTGACTACCGCGGTTTGTAGAGTCTTCTCGATTGTTGT 1453
DB      70482 AGCGTGGAGTTTCCAGGCGATCTGACTATCGTGTATCTATTGTTCTCTAATAATTA 70423
QY      1454 ATTGTTGTTGTGTAATAATTTTATATTTTGAACATCAATTTT 1495
DB      70422 AAAATATATTTTATGCTTTGTATCATTTGTCACAGTAGTTT 70381
RESULT 13
AC134522/c
LOCUS          Medicago truncatula clone mth2-12g4, WORKING DRAFT SEQUENCE.
DEFINITION    AC134522
ACCESSION     AC134522
VERSION       HTG; HTGS_PHASE2; HTGS_DRAFT.
KEYWORDS      Medicago truncatula
SOURCE        Medicago truncatula
ORGANISM      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               rosids; eustersids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
               Medicago.
REFERENCE      1 (bases 1 to 129199)
AUTHORS       Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
               Cook,D., Kim,D. and Roe,B.A.
TITLE         Medicago truncatula BAC Clone mth2-12g4
JOURNAL       Unpublished
REFERENCE      2 (bases 1 to 129199)
AUTHORS       Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
               Cook,D., Kim,D. and Roe,B.A.
TITLE         Direct Submission
JOURNAL       Submitted (27-SEP-2002) Department Of Chemistry And Biochemistry,
               The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
               OK 73019, USA
REFERENCE      3 (bases 1 to 129199)
AUTHORS       Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
               Cook,D., Kim,D. and Roe,B.A.
TITLE         Direct Submission
JOURNAL       Submitted (29-MAR-2003) Department Of Chemistry And Biochemistry,
               The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
               OK 73019, USA
COMMENT       On Mar 29, 2003 this sequence version replaced gi:29336223.
               ----- Genome Center
               Center: Department Of Chemistry And Biochemistry
               The University Of Oklahoma
               Center code:UOKNOR
               -----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes

```

* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 129199: contig of 129199 bp in length.

FEATURES
source
1. 129199
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mth2-1294"
/clone_lib="Medicago truncatula BAC library H2"
BASE COUNT 42930 a 20950 c 21534 g 43771 t 14 others
ORIGIN

Query Match 2.7%; Score 107.6; DB 2; Length 129199;
Best Local Similarity 57.1%; Pred. No. 3.9e-13;
Matches 298; Conservative 0; Mismatches 184; Indels 40; Gaps 4;

Qy 1014 TTCAATGAGCAACAAATGGGACATGATTTGAAGATTCCACGAGACCTCTCGTTATG 1073
Db 70902 TGCAGTATGCTCCTATCGTTGACATTTGAAGATCCCTCGAAACCCACGAGTTATG 70843
Qy 1074 CTTTGTGAGGTAT-----ATTGATCAAGTACAAATTTGTTTTTCTTC 1122
Db 70842 CTTTGTGAGGTATCTACATTTCTTTTGAAGATCGGAAGGGAGTTATCACAC 70783
Qy 1123 TTCTATATGTA-----TAGGCTAATGACTAGATAGTTTGTATGTCGACAGTTGA 1175
Db 70782 TTTTITTAGTCCCTTGTGTTCTCTCTATGATCTTAACATTAATTCAGATTGA 70723
Qy 1176 AGATCCTCGTGATGACAGACATCAATTTATGACGTGATGTTATGATTTGATGGGTG 1235
Db 70722 GGATGCTCGTGATGCTCAAGATCGGATTTATATAGAGATGTTATGATTTGATGTTA 70663
Qy 1236 TCGACTTCGGGTAGTAAACGCAATGAAAGCTAGCTTAATTTCTGTAATTTCTTTGA 1295
Db 70662 TCGATTACTGTTAGATAATTTTCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 70603
Qy 1296 AAGTGTATCTTTGTGTGATGTT-----TTTAGTTGAGATTCGACAT 1339
Db 70602 TCGAATTTTCTCATATGACCTTTTGTGCTACTATTTGTTACTGATGTTGAATTCGACAT 70543
Qy 1340 GGTGTCGTAGATTTTCAACATCAGTTGATAGTACAG-----CAGCAGCTACAGTGG 1393
Db 70542 GGTGACCGGGAAGTTTCATCATCAGTAGACCGATACAGTAGGATACCGTAGAGTGT 70483
Qy 1394 AGCGTGACCTTCAAGACGCTCTGACTACCGGGTTTGTAGAGTCTTCTCGATTGTT 1453
Db 70482 AGCGTGGAGTTTCCAGCGCATCTGACTATCGTGGTATCTATTGTTCTCTAATAATTA 70423
Qy 1454 ATTTGGTGTGTAAATTTTATATTTGAAACTCATTTT 1495
Db 70422 AAAATATATTTTGTGTTGATCATCTGTACAGTAGTTT 70381

RESULT 14
AC012329
LOCUS
DEFINITION Arabidopsis thaliana chromosome 1 BAC T1G12 genomic sequence,
complete sequence.
ACCESSION AC012329
VERSION AC012329.5 GI:12324433
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 80367)
Lin.X., Kaul.S., Town.C.D., Benito.M.-I., Creasy.T.H., Haas.B.J.,
Wu.D., Maiti.R., Ronning.C.M., Koo.H., Fujii.C.Y., Utterback.T.R.,

Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.
Arabidopsis thaliana chromosome 1 BAC T1G12 genomic sequence
2 (bases 1 to 80367)
Lin,X. and Kaul,S.
Direct Submission
Submitted (23-OCT-1999) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, xlin@tigr.org
3 (bases 1 to 80367)
Town,C.D. and Kaul,S.
Direct Submission
Submitted (19-JAN-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
On Jan 19, 2001 this sequence version replaced gi:12280863.
Address all correspondence to:atetigr.org

BAC clone T1G12 is from Arabidopsis thaliana chromosome 1
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of several methods: Gene
prediction programs including Genscan+ (Chris Burge,
http://CCR-081.mit.edu/GENSCAN.html), GenemarkXMM (Mark Borodovsky,
http://genemark.biology.gatech.edu/Genemark/), GlimmerA (a variant
of GlimmerM, see Mihaela Pertea,
http://www.tigr.org/softlab/glimmerm.htm and
GeneSlizer (Mihaela Pertea and Steven Salzberg, contact
mpertea@tigr.org), searches of the complete sequence against a
peptide database and the plant EST database at TIGR
http://www.tigr.org/tdb/tgi.shtml). Annotated genes are named to
indicate the level of evidence for their annotation. Genes with
similarity to other proteins are named after the database hits.
Genes without significant peptide similarity but with EST
similarity are named as unknown proteins. Genes without protein
or EST similarity, that are predicted by more than two gene
prediction programs over most of their length are annotated as
hypothetical proteins. Genes encoding tRNAs are predicted by
tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).
Simple repeats are identified by RepeatMasker (Arian Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html).
Location/Qualifiers
1. 80367
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="1"
/clone="T1G12"
complement(1616..1637)
/rpt_family="AT_rich"
complement(1661..1713)
/rpt_family="(TA)n"
complement(3271..3326)
/rpt_family="AT_rich"
complement(4982..5033)
/rpt_family="AT_rich"
5803..5823
/rpt_family="(A)n"
complement(5873..8780)
/gene="T1G12.1"
/note="contains Pfam profile: PF00083 sugar (and other)
transporter"
complement(join(<5873..5934,6097..6199,6278..6493,
6578..6843,6980..7250,7617..7785,7878..8022,8630..>8780))
/gene="T1G12.1"
complement(join(5873..5934,6097..6199,6278..6493,
6578..6843,6980..7250,7617..7785,7878..8022,8630..8780))
/gene="T1G12.1"
/codon_start=1
/product="putative transporter; 8780-5873"
/protein_id="AAG52174.1"
/db_xref="GI:12324434"
/translation="MEVFYLVFGVMAAVALSKTKNDKINTSSFSFKNNYLL
VFSIMWAGDWLQGPVYVLYSTYFGKGDIGQLFIAGFGSSMLFTGVSLADKQGRK

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

RACVYICIVILSCITRHSPOYKVLVGVGRILGGIATSLIFSAFESWLIHAEHKNRNPQ
QWLSLTSKAVFGLNGVAILGLFGLNLLVDTFSGFVAPFDAAACFLAIGWAILIGT
WSENFQSPDSKDLITQFKVAATAIAADEKXIALGAIQSLFEASMTFVFLWPLALSP
NDEIPIHGFVAFMLASMLGSSLAARMSRLRVENYMQIVFLVSAASLLPIYTS
VLTPSKVDEGLSFIQILGFCVFCVGIWFSIMKORSQYIPEEARSTIMNFF
RVPLNIEVCIVLYNVDAFPIITMFGMCSIFLVASILQRLMLVISEKPAEDWSPMKE
RNSVDLTL

repeat_region 8543..8562
/rpt_family="CAAT)n"
complement(9170..11431)
/gene="TIG12.2"
/note="similar to hypothetical protein (Yer156cp)
GB:6321004 [Saccharomyces cerevisiae]"
complement(9170..9394,9488..9559,9692..9811,
10121..10312,10653..10751,10822..10968,11209..11431))
/gene="TIG12.2"
complement(9170..9394,9488..9559,9692..9811,
10121..10312,10653..10751,10822..10968,11209..11418))
/gene="TIG12.2"
/codon_start=1
/product="unknown protein; 11418-9170"
/protein_id="AAG52181.1"
/db_xref="GI:1232441"
/translation="MFGSRGLCCSRWNRSLFLKRSNFRASPRKRVGTHNGTHCD
EALACFLIRSNRPSDAQIVTRDHOVLEKLDALDVGVDYDQSRKYDHHQKGFSEV
FLGNTKLASAGLVYKYLEILSKLQLEQHPDVFRLFLAYKNFIEAVDALNG
IHQYDTPQPVYNNVSLGHRIGDNLNDWIEPOSSAKDEAFHRAVELASFLSCV
HFHAKSWLPASVMECLAKRYDIDNSGEIKLSKQCPWKLHIFESEEMKIDPPIKY
VLQYDRSENWQIAVSPERFESKALPLAWRLEKEKLSBESSIPRCVFMHSGF
IGANQTVEGALAMARASIMA"

gene 12217..12837
/gene="TIG12.3"
/note="predicted by genscan"
12217..12837
/gene="TIG12.3"
12217..12837
/gene="TIG12.3"
/codon_start=1
/product="hypothetical protein; 12217-12837"
/protein_id="AAG52190.1"
/db_xref="GI:12324450"
/translation="MILPILKSLHIIKTPYEQMAYVHLLMLLIFITPKSSISVVP
DPPTQNLNRIKQITIDFKCNQITISQVLRRTSITKIATKLAIAKAWINAIQTLNLI
EDSLPKARXKRAEDFAMCKRYKLVDAHLNALKYLRDYFRFMYQALALVNLS
MCRYSFFHTPMVYANMKNMLTDIAIYAGKILSPPLTKSKSP"

repeat_region complement(13030..13063)
/rpt_family="AT rich"
complement(14136..15366)
/gene="TIG12.4"
/note="similar to cysteine proteinase GB:AAA50755 [Ainus
glutinosus]; contains Pfam profile: PF00112 papain family
cysteine protease"

gene complement(914136..14716,14922..15366))
/gene="TIG12.4"
complement(914136..14716,14922..15366))
/gene="TIG12.4"
/codon_start=1
/product="putative cysteine proteinase; 15366-14136"
/protein_id="AAG52191.1"
/db_xref="GI:12324451"
/translation="MTSIVFFLLAILLSSRTSGVTSRGLFEASAVEKHEQWMSRFR
VYDSEKTSRFEITNNLKFEVSEINMNTKYTLDVNFEFSLDTEBFKARYTGLVVP
EGMTLITTDSEHTVPSRYENGETGESMDWIOEGATSVKHQOCCGCCWAFSAVA
EGMTKIANGELVLSLEQLLDCSTENNGCGGIMKAPDXIKENOGITEDNVPYOGA
QOTCSNHLAAATISGYETVPCNDEEALLKAGSQVSVSALEGGSYFIHYSGIFNG
ECGTQTHAVTIVGVSEEGIKYWLKNSWGENGYNMRDVSQPMGGLAS
LAYPVA"

repeat_region complement(16903..16941)
/rpt_family="AT rich"
complement(17275..17301)
/rpt_family="AT rich"
complement(18455..20638)
/gene="TIG12.5"

/note="contains pfam profile: PF00566 TBC domain (implies
that these domains are GTPase activator proteins of
Rab-like small GTPases)"
mRNA complement(914136..14716,14922..15366), 19721,19777..19830,19970..20019,
20116..20251,20481..20638))
/gene="TIG12.5"
complement(914136..14716,14922..15366), 19721,19777..19830,19970..20019,
20116..20251,20481..20638))
/gene="TIG12.5"
/codon_start=1
/product="putative GTPase activator protein of Rab-like
small GTPases; 20638-18455"
/protein_id="AAG52193.1"
/db_xref="GI:12324453"
/translation="MKALRIQTSSSNPSSPLSSPPSSSSSSWIIQIRSAFLFVAS
SSPASCSPRLKSPKSRKRLRPOQWKFETPDGLRNGGVLLKVKRSRGIEP
SIRLEVFLGLGLFNGSSKEKERTVNRSSFFDSLAHRCYKCRKEVERLRQCKR
LOKHNGTRKLDNGSETIQDYDMPVDTUSSDESVVSAARESSDEDITDGYM
SEVSTVERDDGSSRRITNATLNSSESSDSDESVYVQFSQATPDVNSTYP
ASSSIPRTEDFVTWQRIRLDAVADSEMTSPSQAVI SEDRACRAEAVGLKDYN
HLEPKYFOAARLVAVLAVLYDPDICYGCGMSDLLSPILSVIPDDHEVFWCFGFM
KKARHFRLEDVGIRQLNIVSKIIKSKDSOLYRHLKLEQADCCFFVYRMVVMFPRR
LTDQTLCLDEVMMADQAIKAGMSKWSIRQAPPTDVLVYALAAASVLQKRRI
IERVSMDEILRECOQSMAGQDVMKLLDDAHLVVTLHTKIEHSFS"

repeat_region complement(23036..23068)
/rpt_family="AT rich"
complement(23851..25075)
/gene="TIG12.6"
/note="similar to 6-phosphogluconolactonase GB:CAB57866
[Homo sapiens]"
complement(23851..24472,24683..24818,24924..25075))
/gene="TIG12.6"
complement(23851..24472,24683..24818,24924..25075))
/gene="TIG12.6"
/codon_start=1
/product="putative 6-phosphogluconolactonase; 25075-23981"
/protein_id="AAG52194.1"
/db_xref="GI:12324454"
/translation="MAPVRRVFKTNEMAVELAKYADLSKFKCKRGVFTVLSGG
DLIAWLKLEAPYIDSIENSKWHIFWDERVCAWDHADSNIKLYADFLSKVPFAE

Query Match 2.6%; Score 105.4; DB 8; Length 80367;
Best Local Similarity 54.0%; Pred. No. 1,2e-12;
Matches 289; Conservative 0; Mismatches 231; Indels 15; Gaps 3;

Qy 717 TTCTTCTTCTTCGAAATATTTTCAGFAATCAATTTCTTCTCTAGATTTTCA 776
Db 53367 TTCTTGTCTATAGCAGTAATTTTTCCTTTATTTTTCCTTTTCTTCA 53426

Qy 777 CAGGAA--CTAATTTCTCTCTGAGTATCAGATGAGTAGCGATGGAATCGTACGATC 834
Db 53427 TGTGCAAGTTTCCCTTTTAACTTTTGAACAATGAGTGGCGGATTTTCTCGGCAATC 53486

Qy 835 TAGCTTGGGAATTTGCTCGAGATATTCGCAAGTGTGAGTTGAAGATCTCTTTACAAAG 894
Db 53487 TAGTTGGTAACCTGCGCGTGACATAGGGAACATGAGATTGAAGATATCTTTTACAAAG 53546

Qy 895 GTTTGAAAATTTCTCTTCTTCGATAAAATGAATTCATTTAGACTAGTTGGGTT 954
Db 53547 GTTCAGTTTGGTATTTTGTTCATTCAGTTACGTCCTCTCTTGAATCGTTATCCCTA 53606

Qy 955 CATAAATTTGCAATTCGTCTGCTGAGACAATTAATTCGACTCTATGATATTTGTT 1014
Db 53607 TAAGACAGTGAAGATTTCTG-----AGCATCTTTTGTATTCATATTAATGCTCTGTTT 53661

Qy 1015 TCAGTATGACCAATTTGCGACATTTGAGATTGACCCAGAGACCTCTGTTGTTATGC 1074
Db 53662 GCAGTATGCCGCAATTTGCGATATTGAATTGAAGGTTCCACCTCGGCTCCCATTTGTTG 53721

Qy 1075 CTTTGTGAGGATATTTGATCAAGTACAAATTTGTTTTTTTCTTCTTCTTGAATAGTAT 1134
Db 53722 CTTTGTGAGGTAAGTATTTGCTTACATTACATGTTCTTCAACCACTTTTAAATTTCCG 53781

exon	4344..4510 /gene="T9C5.10" /number=13 5507..6722 /gene="T9C5.20" join(5507..5537,5770..5861,5962..6042,6163..6260, 6343..6424,6684..6722) /gene="T9C5.20" /note="Gene model changed according to ESTs GB:A1997074 and R65541 similarity to CGI-141 protein- Homo sapiens, EMBL:AF151899" /codon_start=1 /product="putative protein" /protein_id="CAB62447.1" /db_xref="GI:6561943" /translation="MMSFEKNDKIKIGLGTGQVPSFLGIIVFDKGLAMENILF ISGSLIIGKSTQTFPMKRONYKGTISFGVGFVILGWPIGLMLETYGFVFLPSG FWPTLAVFAQKIPVLGIIQPIYIRSFDFKIRGRVDPV" 5507..5537 /gene="T9C5.20" /number=1 5538..5769 /gene="T9C5.20" /number=1 5770..5861 /gene="T9C5.20" /number=2 5862..5961 /gene="T9C5.20" /number=2 5962..6042 /gene="T9C5.20" /number=3 6043..6162 /gene="T9C5.20" /number=3 6163..6260 /gene="T9C5.20" /number=4 6261..6342 /gene="T9C5.20" /number=4 6343..6424 /gene="T9C5.20" /number=5 6425..6683 /gene="T9C5.20" /number=5 6684..6722 /gene="T9C5.20" /number=6 8320..10019 /gene="T9C5.30" join(8320..8406,8526..8591,8670..8744,8899..9040, 9124..9170,9255..9312,9449..9502,9587..9636,9869..10019) /gene="T9C5.30" /note="strong similarity to PRE-MRNA SPLICING FACTOR SF2 - Arabidopsis thaliana, SWISSPROT:SF21_ARATH" /codon_start=1 /product="PRE-MRNA SPLICING FACTOR SF2-like protein" /protein_id="CAB62448.1" /db_xref="GI:6561944" /translation="MSGFRSRIYVGNLPGDIRHEIEDIFYKYGRIVDIELKVPFRP PCYCFVEFHSRDEDAIKGRDGYNLDCRLVELANGRGQSSDRRGVGGGGGY
exon	777 CAGGAA--CTAATTTCTGCTCTGAGGTATCAGATGAGTACGCCGATGGAATCGTACGATC 834 8287 TGTGCAAGTTCCCTTTTAACTTTGCAAAACAATGAGTGGCGGATTTCTCGGCAATC 8346 835 TACGTTGGGAATTCGCTGGAGATATTCGCAAGTGTGAGGTGGAAGATCTCTCTACAAAG 894 8347 TATGTTGGTAATCTCCCGGTGACATTAGGGAACATGAGATTGAAGATATCTTTACAAAG 8406 895 GTTTGAAAAATTTCCCTCTTTCTCTCGATAAAAAATTTGAATTCATTATGACTAGTTTGGGTT 954 8407 GTCAAGTTTGGTTATTTGTTGATTGATTACGGGTCTCTCTTGAATCGTTATCCCTA 8466 955 CATAAATTTGCAATTCGTCTTGTCTGAGACAAATTTAAATTCGACTCTTATGATATTTGTT 1014 8467 TAAGACAGTGAAGATTTCTTG-----AGCATCTTTTGTATTTCATATTAATGCTCTTGT 8521 1015 TCAGTATCGACCAATTTGGACATTTGGAAGATTCCACGAGACCTCTCTGGTTATGC 1074 8522 GCAGTATGCCCGCATTTGCGATATTGAATTTGAAGTTCCACCTCGGCTCCATGTTATG 8591 1075 CTTTGTGAGGTATATTGATCAAGTACAAATTTGTTTTTTTCTTCTTGTGTAATAGTAT 1134 8582 CTTTGTGAGGTAAAGTATTTCGCTTACATTACATGTTCCCTAACCACTTTTAAATTCGC 8641 1135 AGGCTAATGACTAAGATAGTTTGTATTGTTGGTGGCAGTTTGAAGATCCCTCGTATGCAGAC 1194 8642 TGAACCAAGCTC-----TGTTGAATGTTTGTAGTTTGACATTTCTCGGATCTGAA 8693 1195 GATGCAATTTATGGACGTGATGTTATGATTTGATGGGTGTCGACTTCGGGTTA 1249 8694 GATGCCATCAGGGCGCGTATGCGCTATAATTTGGATGGCTGTCTGCTGAGGGTAA 8748

Search completed: January 28, 2004, 21:00:52
Job time : 14437 secs

Query Match	2.6%; Score 105.4; DB 8; Length 104204;
Best Local Similarity	54.0%; Pred. No. 1.2e-12;
Matches 289; Conservative	0; Mismatches 231; Indels 15; Gaps 3;
QY 717	TTCTTCTCTCTCGAATTAATTTTCAGTAATCAATTTCTCTCTCTCTAGATTTTTA 776
DB 8227	TTCTTGTGCTATAGCAGTAATTTGATTCCTTTATTTATTCGTTCTCTTTTATCTCA 8286

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 28, 2004, 16:22:45 ; Search time 1239 Seconds
(without alignments)

11673.323 Million cell updates/sec

Title: US-10-014-927-18

Perfect score: 4044

Sequence: 1 agacaagatgcttactctct.....atgattatcttaactaaac 4044

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2356869 seqs, 1788235258 residues

Total number of hits satisfying chosen parameters: 4713738

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4044	100.0	4044	14	US-10-014-927-18
2	780	19.3	780	10	US-09-938-842A-3893
3	780	19.3	780	12	US-09-938-842A-3893
4	108.8	2.7	762	10	US-09-938-842A-1199
5	108.8	2.7	762	12	US-09-938-842A-1199
6	65.4	1.6	6161	13	US-10-311-455-384
7	64.8	1.6	14006	13	US-10-311-455-1931
8	64.4	1.6	6171	13	US-10-311-455-761
9	63.2	1.6	6668	13	US-10-311-455-1670
10	62.6	1.5	3673778	13	US-10-312-841-2
11	62	1.5	1267	14	US-10-001-843-45
12	61.4	1.5	529	10	US-09-983-965-2109
13	61.2	1.5	6334	13	US-10-311-455-1186
14	59.6	1.5	6668	13	US-10-311-455-1669
15	59	1.5	6338	13	US-10-311-455-1099

16	58.4	1.4	9539	13	US-10-240-453-53	Sequence 53, Appl
17	58.4	1.4	9539	15	US-10-239-676-51	Sequence 51, Appl
18	58.2	1.4	16033	13	US-10-311-455-1377	Sequence 1377, Ap
19	58.2	1.4	17934	13	US-10-311-455-1692	Sequence 1692, Ap
20	58	1.4	9539	13	US-10-240-453-54	Sequence 54, Appl
21	58	1.4	9539	15	US-10-239-676-52	Sequence 52, Appl
22	57.8	1.4	6621	15	US-10-172-086-9	Sequence 9, Appl
23	57.4	1.4	7571	13	US-10-311-455-500	Sequence 500, App
24	56.6	1.4	446	10	US-09-960-352-3400	Sequence 3400, Ap
25	56	1.4	431	10	US-09-960-352-4617	Sequence 4617, Ap
c	56	1.4	629	13	US-10-125-968-78	Sequence 78, Appl
27	55.8	1.4	1223197	13	US-10-027-632-179264	Sequence 179264,
28	55.8	1.4	1223197	14	US-10-027-632-179264	Sequence 179264,
29	55.2	1.4	2140405	13	US-10-027-632-76212	Sequence 76212, A
30	55.2	1.4	2140405	14	US-10-027-632-76212	Sequence 76212, A
31	54	1.3	439	10	US-09-960-352-1009	Sequence 1009, Ap
32	54	1.3	7657	13	US-10-311-455-1995	Sequence 1995, Ap
33	54	1.3	7657	15	US-10-239-676-185	Sequence 185, App
34	54	1.3	12007	13	US-10-311-455-690	Sequence 690, App
35	53.6	1.3	5822	13	US-10-311-455-1069	Sequence 1069, Ap
36	53.6	1.3	6145	13	US-10-311-455-945	Sequence 945, App
37	53.4	1.3	345	9	US-09-770-791-783	Sequence 783, App
38	53.4	1.3	11996	13	US-10-240-485-45	Sequence 45, Appl
39	53.4	1.3	113515	13	US-10-311-455-2147	Sequence 2147, Ap
40	52.8	1.3	6485	13	US-10-311-455-1781	Sequence 1781, Ap
41	52.8	1.3	7057	13	US-10-311-455-1822	Sequence 1822, Ap
42	52.8	1.3	7057	13	US-10-240-485-148	Sequence 148, App
43	52.8	1.3	12138	13	US-10-311-455-1601	Sequence 1601, Ap
44	52.8	1.3	3673778	13	US-10-312-841-1	Sequence 1, Appl
c	52.6	1.3	525	15	US-10-198-846-1483	Sequence 1483, Ap

ALIGNMENTS

RESULT 1
US-10-014-927-18
; Sequence 18, Application US/10014927
; Publication No. US20020115180A1
; GENERAL INFORMATION:
; APPLICANT: Barta, Andrea
; APPLICANT: Lopato, Sergyi
; APPLICANT: Kalyna, Maria
; APPLICANT: Dörner, Silke
; TITLE OF INVENTION: Splice Factor
; FILE REFERENCE: SONN:013US
; CURRENT APPLICATION NUMBER: US/10/014,927
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: PCT/AT00/00100
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: A 727/99
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 4044
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of the unknown organism: genome
; OTHER INFORMATION: atSRp30
US-10-014-927-18

Query Match	100.0%;	Score 4044;	DB 14;	Length 4044;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 4044;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AGACAAAGATGCTTACTCTTAAACATGTTTCGAGGTTTATTGAAATGATCACCAGCTTC	60	
Db	1	AGACAAAGATGCTTACTCTTAAACATGTTTCGAGGTTTATTGAAATGATCACCAGCTTC	60	
Qy	61	TAACTATGGTATCTTCTTCCTCGAAGCGAAGATGATTGAACCGAAATAT	120	

Db 61 TAACTATGGTATCTTCTCCCTGCAAGCAAGTGGAGATGATTGATATAACGGAATAT 120
Qy 121 CGGAACATCACTCAACAAACCAAAATTTGGACATCATATCGCAACAAATTTCAATAGGAA 180
Db 121 CGGAACATCACTCAACAAACCAAAATTTGGACATCATATCGCAACAAATTTCAATAGGAA 180
Qy 181 AAATACTGAAATTCCTCAACAAACCAAAATTTGGACATCATATCGCAACAAATTTCAATAGGAA 240
Db 181 AAATACTGAAATTCCTCAACAAACCAAAATTTGGACATCATATCGCAACAAATTTCAATAGGAA 240
Qy 241 GAGAGACCGTGGACCGTGTCAACCGCGGAAATGGTCTTTGACGAGAGTTACTAATCGGCGA 300
Db 241 GAGAGACCGTGGACCGTGTCAACCGCGGAAATGGTCTTTGACGAGAGTTACTAATCGGCGA 300
Qy 301 ATTGAGATTTGAGAGTGGTAGTAGAGAAACCGAGAGAAATGTTTCTCTCAAAAAATATCCC 360
Db 301 ATTGAGATTTGAGAGTGGTAGTAGAGAAACCGAGAGAAATGTTTCTCTCAAAAAATATCCC 360
Qy 361 CAAGTGTCTCCGATCTAGTGTCTCTTTGTCGAAACCGAGAGATGTTTCTCTCAAAAAATATCCC 420
Db 361 CAAGTGTCTCCGATCTAGTGTCTCTTTGTCGAAACCGAGAGATGTTTCTCTCAAAAAATATCCC 420
Qy 421 AGAATGAATGACCGATGGTCCGAAATCGGATTCGAAATTTGGTTCGAAATTTGTAATACTAA 480
Db 421 AGAATGAATGACCGATGGTCCGAAATCGGATTCGAAATTTGGTTCGAAATTTGTAATACTAA 480
Qy 481 CATACAAATATTTCCGGTTTGAATGATAAGAAAAACACATTCGATCCGGTTAGAACAAATAT 540
Db 481 CATACAAATATTTCCGGTTTGAATGATAAGAAAAACACATTCGATCCGGTTAGAACAAATAT 540
Qy 541 ATTAACAGGCCATTAATAACATATGGGCCGATCTTGATCAACTGGGCTATTTCATCGTTGA 600
Db 541 ATTAACAGGCCATTAATAACATATGGGCCGATCTTGATCAACTGGGCTATTTCATCGTTGA 600
Qy 601 TACATGGCGCGCACAGATTAATAACATATGGGCCGATCTTGATCAACTGGGCTATTTCATCGTTGA 660
Db 601 TACATGGCGCGCACAGATTAATAACATATGGGCCGATCTTGATCAACTGGGCTATTTCATCGTTGA 660
Qy 661 AACGAAACGGTGGTGTCTCTCCAGAACAAATCTAACGGTCTTCTCGAAACATCTTCTCT 720
Db 661 AACGAAACGGTGGTGTCTCTCCAGAACAAATCTAACGGTCTTCTCGAAACATCTTCTCT 720
Qy 721 TCTTCTCTCTCGAAATATTTTTCAGTAATCAATTTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Db 721 TCTTCTCTCTCGAAATATTTTTCAGTAATCAATTTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Qy 781 AACTAAATTTCT 840
Db 781 AACTAAATTTCT 840
Qy 841 GGGAAATTTGCTGAGATATTCGCAAGTGTGAGGTTGCAAGATCTCTCTCTCAAGGTTTGA 900
Db 841 GGGAAATTTGCTGAGATATTCGCAAGTGTGAGGTTGCAAGATCTCTCTCTCAAGGTTTGA 900
Qy 901 AAAATTTCT 960
Db 901 AAAATTTCT 960
Qy 961 TTTGCAATTTCT 1020
Db 961 TTTGCAATTTCT 1020
Qy 1021 TGGACCAATTTGAGACATTTGAGATTTCCACCGACCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
Db 1021 TGGACCAATTTGAGACATTTGAGATTTCCACCGACCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
Qy 1081 CGAGGTATTTGAGTCAAGTACAAATTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
Db 1081 CGAGGTATTTGAGTCAAGTACAAATTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
Qy 1141 ATGACTAAGATAGTTTGTATTTGTTGCGGAGTTTGAAGATCTCTCTCTCTCTCTCTCTCTCT 1200
Db 1141 ATGACTAAGATAGTTTGTATTTGTTGCGGAGTTTGAAGATCTCTCTCTCTCTCTCTCTCTCT 1200

Qy 1201 ATTTATGACGCTGATGGTTATGATTTTGGTGGTGCAGCTTCGGGTTAGTAAACGCGATG 1260
Db 1201 ATTTATGACGCTGATGGTTATGATTTTGGTGGTGCAGCTTCGGGTTAGTAAACGCGATG 1260
Qy 1261 ATGAAAGCTAGCTTAAATTTTCTGTAATTTCTGTAAGGTTTATCTTTGTTGTTGATGTTT 1320
Db 1261 ATGAAAGCTAGCTTAAATTTTCTGTAATTTCTGTAAGGTTTATCTTTTGTGTTGATGTTT 1320
Qy 1321 TTAGGTTGAGATTTGACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1380
Db 1321 TTAGGTTGAGATTTGACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1380
Qy 1381 CAGCTTACAGTCCGAGCGCTGACCTTCAAGACGCTCTGACTACCGCGGTTTGTAGAGTCT 1440
Db 1381 CAGCTTACAGTCCGAGCGCTGACCTTCAAGACGCTCTGACTACCGCGGTTTGTAGAGTCT 1440
Qy 1441 TCTCGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1500
Db 1441 TCTCGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1500
Qy 1501 CTAAACATGTTAGTCTTGTGACCGGATTTACCGCTTCTGCTTCTGCTGCGCAGGACCTTAAG 1560
Db 1501 CTAAACATGTTAGTCTTGTGACCGGATTTACCGCTTCTGCTTCTGCTGCGCAGGACCTTAAG 1560
Qy 1561 TAAAGGACACTATATAGTCTTTTCTCTGTAATGTTGTTCTCTATATCATGTTTGGATT 1620
Db 1561 TAAAGGACACTATATAGTCTTTTCTCTGTAATGTTGTTCTCTATATCATGTTTGGATT 1620
Qy 1621 TATCTCTTTCTGAAATGATGTTATTTGCTATTTTACCGGTTGATTTAGGATCATCGGCAA 1680
Db 1621 TATCTCTTTCTGAAATGATGTTATTTGCTATTTTACCGGTTGATTTAGGATCATCGGCAA 1680
Qy 1681 AGCTGAGATGCTGCTTCTCTGAAATGTTTCCCTGACCGTAAAGGTGAGTTGACATTCGA 1740
Db 1681 AGCTGAGATGCTGCTTCTCTGAAATGTTTCCCTGACCGTAAAGGTGAGTTGACATTCGA 1740
Qy 1741 TAGTTTGGATTAAGCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1800
Db 1741 TAGTTTGGATTAAGCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1800
Qy 1801 GTGTTAAGCACTGAACTGCTAAACCTCACATTTCTGTTAGGATTTCTTTGTTAGGCAATG 1860
Db 1801 GTGTTAAGCACTGAACTGCTAAACCTCACATTTCTGTTAGGATTTCTTTGTTAGGCAATG 1860
Qy 1861 TTGTTGATTTATAGCAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
Db 1861 TTGTTGATTTATAGCAATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
Qy 1921 GCAAAATGTTCTGGAATTTATGCTTTGATGCTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1980
Db 1921 GCAAAATGTTCTGGAATTTATGCTTTGATGCTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1980
Qy 1981 TTATCTCTCTGAGCGCAGTGTGATGCTTTGATGCTTTGATGCTTTGATGCTTTGATGCTGAC 2040
Db 1981 TTATCTCTCTGAGCGCAGTGTGATGCTTTGATGCTTTGATGCTTTGATGCTTTGATGCTGAC 2040
Qy 2041 GATAAGGAACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
Db 2041 GATAAGGAACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
Qy 2101 ATGTTGATTTGCTTTCTTTGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 2160
Db 2101 ATGTTGATTTGCTTTCTTTGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 2160
Qy 2161 TTACTGTTTCAAGTGGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
Db 2161 TTACTGTTTCAAGTGGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
Qy 2221 AAAGCTATAGAGCAGGAGTCCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
Db 2221 AAAGCTATAGAGCAGGAGTCCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280

US-09-938-842A-3893

Query Match 19.3%; Score 780; DB 10; Length 780;
Best Local Similarity 100.0%; Pred. No. 3.8e-187;
Matches 780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	28	GTTCGAGGTTTATTGAAATGATCACCAGCTTCTAATCTGATGTTCTTCCCTGCAAG	87
Db	1	GTTCGAGGTTTATTGAAATGATCACCAGCTTCTAATCTGATGTTCTTCCCTGCAAG	60
Qy	88	CGAACAGTGAAGATGATTGATAACCGAAATATCGGAACATCACTCAACAAACCAAAAT	147
Db	61	CGAACAGTGAAGATGATTGATAACCGAAATATCGGAACATCACTCAACAAACCAAAAT	120
Qy	148	TTGGACATCATATCGCAACAAATTCATAGGAAATATCTGAAATTCGAAACAGAAAAA	207
Db	121	TTGGACATCATATCGCAACAAATTCATAGGAAATATCTGAAATTCGAAACAGAAAAA	180
Qy	208	CGAAACGGAACAGAGCAGGAACTCAGGACTGAGAGAGCCGTGGACGGTGTCAAGGGG	267
Db	181	CGAAACGGAACAGAGCAGGAACTCAGGACTGAGAGAGCCGTGGACGGTGTCAAGGGG	240
Qy	268	AAATGCTTGTGACGGAGTTACTTAATCGCGAAATTCGAAATTCGAAATTCGAAATTC	327
Db	241	AAATGCTTGTGACGGAGTTACTTAATCGCGAAATTCGAAATTCGAAATTCGAAATTC	300
Qy	328	GAACCCGAGAGATGTTTCTCTCAAAAAATCCCAAGTGTTCGGAATCTAGTGTCTCTTT	387
Db	301	GAACCCGAGAGATGTTTCTCTCAAAAAATCCCAAGTGTTCGGAATCTAGTGTCTCTTT	360
Qy	388	TGTCGCAAAAACGACAGTGTTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTC	447
Db	361	TGTCGCAAAAACGACAGTGTTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTC	420
Qy	448	CGATCTTGATCAACTGGGCTATTCATCGTGTGATACATCGGCGGCACAGGATTAATC	507
Db	421	CGATCTTGATCAACTGGGCTATTCATCGTGTGATACATCGGCGGCACAGGATTAATC	480
Qy	568	CAGTTCGGTTTATAAAGATGATCTAGTTTCGAAACGAGCGGTGTCTCTCTCCAG	627
Db	541	CAGTTCGGTTTATAAAGATGATCTAGTTTCGAAACGAGCGGTGTCTCTCTCCAG	600
Qy	628	AACAATCTAAGCTTCTCGAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	687
Db	601	AACAATCTAAGCTTCTCGAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	660
Qy	748	TAATCAATTTCT	807
Db	721	TAATCAATTTCT	780

RESULT 3
US-09-938-842A-3893
; Sequence 3893, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krebs, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938.842A
; CURRENT FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 3893
; LENGTH: 780
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-3893

Query Match 19.3%; Score 780; DB 12; Length 780;
Best Local Similarity 100.0%; Pred. No. 3.8e-187;
Matches 780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	28	GTTCGAGGTTTATTGAAATGATCACCAGCTTCTAATCTGATGTTCTTCCCTGCAAG	87
Db	1	GTTCGAGGTTTATTGAAATGATCACCAGCTTCTAATCTGATGTTCTTCCCTGCAAG	60
Qy	88	CGAACAGTGAAGATGATTGATAACCGAAATATCGGAACATCACTCAACAAACCAAAAT	147
Db	61	CGAACAGTGAAGATGATTGATAACCGAAATATCGGAACATCACTCAACAAACCAAAAT	120
Qy	148	TTGGACATCATATCGCAACAAATTCATAGGAAATATCTGAAATTCGAAACAGAAAAA	207
Db	121	TTGGACATCATATCGCAACAAATTCATAGGAAATATCTGAAATTCGAAACAGAAAAA	180
Qy	208	CGAAACGGAACAGAGCAGGAACTCAGGACTGAGAGAGCCGTGGACGGTGTCAAGGGG	267
Db	181	CGAAACGGAACAGAGCAGGAACTCAGGACTGAGAGAGCCGTGGACGGTGTCAAGGGG	240
Qy	268	AAATGCTTGTGACGGAGTTACTTAATCGCGAAATTCGAAATTCGAAATTCGAAATTC	327
Db	241	AAATGCTTGTGACGGAGTTACTTAATCGCGAAATTCGAAATTCGAAATTCGAAATTC	300
Qy	328	GAACCCGAGAGATGTTTCTCTCAAAAAATCCCAAGTGTTCGGAATCTAGTGTCTCTTT	387
Db	301	GAACCCGAGAGATGTTTCTCTCAAAAAATCCCAAGTGTTCGGAATCTAGTGTCTCTTT	360
Qy	388	TGTCGCAAAAACGACAGTGTTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTC	447
Db	361	TGTCGCAAAAACGACAGTGTTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTC	420
Qy	448	CGATCTTGATCAACTGGGCTATTCATCGTGTGATACATCGGCGGCACAGGATTAATC	507
Db	421	CGATCTTGATCAACTGGGCTATTCATCGTGTGATACATCGGCGGCACAGGATTAATC	480
Qy	568	CAGTTCGGTTTATAAAGATGATCTAGTTTCGAAACGAGCGGTGTCTCTCTCCAG	627
Db	541	CAGTTCGGTTTATAAAGATGATCTAGTTTCGAAACGAGCGGTGTCTCTCTCCAG	600
Qy	628	CAGTTCGGTTTATAAAGATGATCTAGTTTCGAAACGAGCGGTGTCTCTCTCCAG	687
Db	601	CAGTTCGGTTTATAAAGATGATCTAGTTTCGAAACGAGCGGTGTCTCTCTCCAG	660
Qy	748	TAATCAATTTCT	807
Db	721	TAATCAATTTCT	780

RESULT 4
US-09-938-842A-1199
; Sequence 1199, Application US/09938842A

```

; Patent No. US20020160378A1
;
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
;
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
;
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
;
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
;
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
;
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
;
; NUMBER OF SEQ ID NOS: 5379
;
; SEQ ID NO 1199
; LENGTH: 762
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-938-842A-1199

```

	Query Match	2.7%;	Score 108.8;	DB 10;	Length 762;
	Best local Similarity	90.6%;	Pred. No. 8.9e-17;		
	Matches 116;	Conservative 0;	Mismatches 12;	Indels 0;	Gaps 0;
Qy	1300	TGTTATCTTTGTGTGATGTTT	TAGGTTGAGATTGCACATGGT	TGGTGGTGTAGATTTTCACC	1359
Db	204	TGATTTTGATGGGTGTCAC	CTTCGGGTGAGATTGCACATGGT	TGGTGGTGTAGATTTTCACC	263
Qy	1360	ATCAGTTGATAGGTACAGCAG	CAGCTACAGTGCAGCGCGTGC	ACCTTCAAGACGCTCTGA	1419
Db	264	ATCAGTTGATAGGTACAGCAG	CAGCTACAGTGCAGCGCGTGC	ACCTTCAAGACGCTCTGA	323
Qy	1420	CTACCGCG	1427		
Db	324	CTACCGCG	331		

RESULT 5
US-09-938-842A-1199
; Sequence 1199, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1199
; LENGTH: 762
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1199

Query Match 2.7%; Score 108.8; DB 12; Length 762;
Best Local Similarity 90.6%; Pred. No. 8.9e-17;
Matches 116; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```

Db      204  TGATTTTGATGGGTGTGCAGCTTCGGGTGTGAGATTGCACATGGTGGTGGTATGATTTTCACC 263
Oy      1360  ATCAGTTGATAGGTACAGCAGCAGCTACAGTCAGTCGAGCGGTGCACCTTCAAGACGGCTCTGA 1419
Db      264  ATCAGTTGATAGGTACAGCAGCAGCTACAGTCGAGCGGTGCACCTTCAAGACGGCTCTGA 323
Oy      1420  CTACCGCG 1427
Db      324  CTACCGCG 331

RESULT 6
US-10-311-455-384
; Sequence 384, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 384
; LENGTH: 6161
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-384

```

Query Match	1.6%	Score 65.4	DB 13	Length 6161
Best Local Similarity	47.1%	Pred. No. 3.6e-05		
Matches 201	Conservative 0	Mismatches 226	Indels 0	Gaps 0
Qy	2455	TC AAGGTATCAGGTGTAGATTTCGTATCATTTATATATATAGTAGTTACCCCTTCATGGATC	2514	
Db	2733	TGAAATTTGCTATTATTAAATTTTAAAGTGATATAATTAGGGGTATTTAGTATGTTTA	2792	
Qy	2515	ACTTGTTCCTGGCATAGTGAACCTCCTACTACGTCTTTATTTACTTACAACTAAGCACTTTTG	2574	
Db	2793	TAGTGTATATGTAGTATGAAATTTAAATTTTGAAGAATATTTTATTTATTTTAAATAGAGTIT	2852	
Qy	2575	TTGCTTCCGTACACAGTGTGAATTTGTTGAGTCTTTTTTCCCTCATAGTGGACTAGTCTA	2634	
Db	2853	TATATTTAAATAAGTAGTATTATTATTTTGTTTTTTTTTTAGTTTTTGGTAATATAAAT	2912	
Qy	2635	TTGTCACTTGATTTTCTCTCTTGTTGATGTTTCTCATGTCATGCAAACTCCCAATATGGG	2694	
Db	2913	GTGTTTTTGTGTTTTATGGAATTTGTTTTTATTTTGAATATTTTAAGTAAATAGAGTTTATATA	2972	
Qy	2595	TAAAGGTACCTCCTCTGTTGGGATTACCAGAGTTCCTATTCTCATTTCTACACGTGAATG	2754	
Db	2973	GTGTGTGTGATATTTTCGTTTTTCGGGTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	3032	
Qy	2755	TGTTTGTTTTTTATGTTTTTGAGTCTTTGACAGAGATGCTCCCATCATATTTAGTCCTTTT	2814	
Db	3033	TT	3092	
Qy	2815	CCCTTCTCTCTGTGTGGTCTCTCTCTGGATGTTCCCTCTGTGATAAAGCTTTACTTCTTAA	2874	
Db	3093	TT	3152	
Qy	2875	CTTTTTT 2881		

Db 3153 TTTT 3159

RESULT 7
US-10-311-455-1931
; Sequence 1931, Application US/10311455
; Publication No. US20030143608A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1931
; LENGTH: 14006
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: 8289, 8310, 8313
; OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1931

Query Match 1.6%; Score 64.8; DB 13; Length 14006;
Best Local Similarity 44.3%; Pred. No. 8.8e-05;
Matches 264; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

QY	2286	TC	TTTTTTTTTTTTTTTTTTTTTTTTCATAAACCTTAAGACATATAAGGATTTTTTATTGTAAC	TT	2345
Db <td>1892</td> <td>TTTTGG</td> <td>TTTTTTTTTTTTTTTTTTTTTTTATTTT</td> <td>TTTTTTTTTTTTTTTTTTTTTTTTTTTTTT</td> <td>1951</td>	1892	TTTTGG	TTTTTTTTTTTTTTTTTTTTTTTATTTT	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	1951
QY <td>2346</td> <td>ATTATGAAATTA</td> <td>ACTGACITCTCTAAATGCAATGAGTGTCACCTCTAGATCCATTTC</td> <td></td> <td>2405</td>	2346	ATTATGAAATTA	ACTGACITCTCTAAATGCAATGAGTGTCACCTCTAGATCCATTTC		2405
Db <td>1952</td> <td>TT</td> <td></td> <td>TT</td> <td>2011</td>	1952	TT		TT	2011
QY <td>2406</td> <td>CCCGGTTACAGGCC</td> <td>CCCTTAGTCGTCCTCGCTATACAGCTCTGCTCAAGGTATGA</td> <td></td> <td>2465</td>	2406	CCCGGTTACAGGCC	CCCTTAGTCGTCCTCGCTATACAGCTCTGCTCAAGGTATGA		2465
Db <td>2012</td> <td>TT</td> <td></td> <td>TT</td> <td>2071</td>	2012	TT		TT	2071
QY <td>2466</td> <td>GTGTAGATTTGATCAT</td> <td>TATATATAGTATGATACCCCTTCATGGATCACTTGTTCTTG</td> <td></td> <td>2525</td>	2466	GTGTAGATTTGATCAT	TATATATAGTATGATACCCCTTCATGGATCACTTGTTCTTG		2525
Db <td>2072</td> <td>TT</td> <td></td> <td>TT</td> <td>2131</td>	2072	TT		TT	2131
QY <td>2526</td> <td>CATAGTGA</td> <td>CTCCTTACTAGCTTTTATTTACTTTACAACTAAGCACCTTTTGTTGCTTCGTA</td> <td></td> <td>2585</td>	2526	CATAGTGA	CTCCTTACTAGCTTTTATTTACTTTACAACTAAGCACCTTTTGTTGCTTCGTA		2585
Db <td>2132</td> <td>TTTTTTTTTTTTTTTTTTTTTTTGA</td> <td>TT</td> <td></td> <td>2191</td>	2132	TTTTTTTTTTTTTTTTTTTTTTTGA	TT		2191
QY <td>2586</td> <td>CACAGTTGA</td> <td>ATTTGTTTGGAGTCITTTTTCCCTCATAGTGGACTAGTCTATTGCACTGA</td> <td></td> <td>2645</td>	2586	CACAGTTGA	ATTTGTTTGGAGTCITTTTTCCCTCATAGTGGACTAGTCTATTGCACTGA		2645
Db <td>2192</td> <td>TT</td> <td></td> <td>TT</td> <td>2251</td>	2192	TT		TT	2251
QY <td>2646</td> <td>TTTTTCTCCTTTGTTGA</td> <td>TTTTTCTATGTGTCATGTCATGCAAACTCAATATGGGTAAAGGTACC</td> <td></td> <td>2705</td>	2646	TTTTTCTCCTTTGTTGA	TTTTTCTATGTGTCATGTCATGCAAACTCAATATGGGTAAAGGTACC		2705
Db <td>2252</td> <td>TT</td> <td></td> <td>TT</td> <td>2311</td>	2252	TT		TT	2311
QY <td>2706</td> <td>TCCTTGT</td> <td>TTGGATTAACAGAGTTCCCTTTTTCATTTCTTACACGTGAAGGTTGTTTTT</td> <td></td> <td>2765</td>	2706	TCCTTGT	TTGGATTAACAGAGTTCCCTTTTTCATTTCTTACACGTGAAGGTTGTTTTT		2765
Db <td>2312</td> <td>TT</td> <td></td> <td>TT</td> <td>2371</td>	2312	TT		TT	2371
QY <td>2766</td> <td>TATGTTT</td> <td>TGATGTTCTTGACAGAGATGCTCCCATCATATTTAGTCCCTTTTTCCTTCTCTTT</td> <td></td> <td>2825</td>	2766	TATGTTT	TGATGTTCTTGACAGAGATGCTCCCATCATATTTAGTCCCTTTTTCCTTCTCTTT		2825

[illegible]

Query Match	1.6%	Score 64.8	DB 13	Length 14006
Best Local Similarity	44.3%	Pred. No. 8.8e-05		
Matches 264	Conservative 0	Mismatches 332	Indels 0	Gaps 0
QY	2286	TCCTTTTTTTTTTTTTTTTCATTAACCTAAGACATATAAGGATATTTTATCTACATT	2345	
Db	1992	TTTTGGTTT	1951	
QY	2346	ATTATGAAATTAACAGTCACTCTCTAAATGCAATGCAGTGTGTCACTGCTAGATCCATTTC	2405	
Db	1952	TTTTTTTTTTATTGATTT	2011	
QY	2406	CCCGGTTACGGCCCTTAGTCTCTCGCTACAGCTCTGTCTCAAGGTATGA	2465	
Db	2012	TTT	2071	
QY	2466	GTGTTAGATTGTATCAATTATATATATGATGATTTACCCCTTCATGGATCACTGTGCTTG	2525	
Db	2072	TTT	2131	
QY	2526	CATAGTGAACCTCTACTAGCTTTATTACTTACAACTAAGCACCTTTGTTGCTTCGTA	2585	
Db	2132	TTTTTTTTTTTTTTTTTTCGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	2191	
QY	2586	CACAGTTGAATTTGTTGAGTCTTTTTTCCCTCATAGTGGACTAGTCTATTGTCACATGA	2645	
Db	2192	TTTTTTTTTTATTCGTTTTT	2251	
QY	2646	TTTTTCTCCTTTGTTGATGTTTTCTATGTCTATGTCAAGAACTCAATATCGGTAAGGTTACC	2705	
Db	2252	TTTTTTTTTTATTCGTTTTTTT	2311	
QY	2706	TCCCTGTTTGGGATTAACAGAGTTCCTTTTTCATTTCTTACACGTGAATGTTGTTTTT	2765	
Db	2312	TTTTTTTTTTTTTTATATT	2371	
QY	2766	TATGTTTTGATGTTCTTGACAGAGATGCTCCCATCATATTTAGTGCCTTTTCTCTTCT	2825	

	Query Match	1.6%	Score 64.4	DB 13	Length 6171
	Best Local Similarity	46.1%	Pred. No. 6.5e-05		
	Matches 215	Conservative 0	Mismatches 251	Indels 0	Gaps 0
Qy	2433	TCGCTGCCTATACAGCTCTGCTCAAGGTATGAGTGTAGATTGTGATCATATTATATATA	2492		
Db	189	TAGTTGGGATTATAGATTAAGTTATTATATTGGTTAAATTTTGTATTTTGTAGTAGAGA	248		
Qy	2493	TGTAGTTACCCCTTCATGGATCACTTGTTCTTCGCATAGTGAACCTCCTTACTAGCTTTATT	2552		
Db	249	CGGGTTTATTATGATGTTGGTTAGTTGGTTTAAATTTTGATTTCTGTGATTTTTTTGTT	308		
Qy	2553	ACTTACAACATAAGCAGCCTTTTGTGCTTCGGTACACAGATGAATTTGTTGAGCTCTTTT	2612		
Db	309	TTAGTTTTTAAATTGTTGGATTATAGGCGTGAGTTTATTTTTCGGATTTTTTTTTT	368		
Qy	2613	TCCCTCATAGTGACATAGTCTATTGTGTCATCTTGATTTTCTTCTTGTGTGATTTCTAT	2672		
Db	369	TT	428		
Qy	2673	GTCAAGCAAACTCCATAAGGGTAAAGGTTACCTCCTGTTGGGATTACCAGAGTTCC	2732		
Db	429	TT	488		
Qy	2733	TTTCATTCTTTACACGTGAATGTTTGTGTTTTTATGTTTTGAGTTCTTGACAGATGCG	2792		
Db	489	TT	548		
Qy	2793	TCCCATCATATTAGTGCCTTTTCTTTCTTTGTGTCGTTCTCTCTCTGATGTTTTCC	2852		
Db	549	TT	608		
Qy	2853	CTGATAAAGCTTTACTCTTAACTTTTTTCCAGCGACGGTGAATTT	2998		
Db	609	TTGAGATAGGGTTTTATTGTTGTTTAAAGTAAACGTGATTTTATTT	654		

654

[illegible][illegible]

RESULT 11
US-10-001-843-45/c
; Sequence 45, Application US/10001843
; Publication No. US902013225A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Caffery, Robert
; APPLICANT: Sun, Yongming

RESULT 12
US-09-983-965-2109
; Sequence 2109, Application US/0983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678

RESULT 13
US-10-311-455-1186
; Sequence 1186, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1186
; LENGTH: 6334
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1186

Query Match 1.5%; Score 61.2; DB 13; Length 6334;
Best Local Similarity 44.9%; Pred. No. 0.00043;
Matches 273; Conservative 0; Mismatches 333; Indels 2; Gaps 1;
QY 2449 TCTGCTCAAGTATGAGTATGATTTCTATCATTATATATATATAGTAGTACCCCTTCA 2508
Db 364 TATTTTCGTTTTTCGGGTTTTTTTCAGGTTTTTTTTTTTTTTTGGGTTTTTTTTTAA 423
QY 2509 TGGATCAGTTGTTCTTGATAGTGAACCTCTTACTAGCTTTATTTACTTACAACTAAGCAC 2568
Db 424 TTTTTCGTTTTTTTCGGGTTTTTTTCGTTTATCGGTTTTTTTTTTTTTTTCGGGTT 483
QY 2569 CTTTGTGCTCCGACACAGTGAATTTGTTGAGTCTTTTCCCTCATAGTGGACT 2628
Db 484 TTTTTCGTTTTTCG-GGTTTTTTTTTTTTTTCGAAATTTTTTTTTTTTTTTTCGAGTTTT 541
QY 2629 AGTCATATCTCACTGATTTTCTCTTTGTTGATGTTTTCTATCATGCAAACTCCAA 2688
Db 542 TTTTTCGTTTTTCGAGTTTTTTTTTTTTTTTAAATTTTTTTTTTTTCGGGTTTTTTT 601
QY 2689 TATGGTAAGGTACTCTCTGTTTGGATTACAGAGTCTCTTTTCATTTCTTTACACG 2748
Db 602 TTTTTCGAGTTTTTTTATTTATATGCGTTTTTTTGTGTTTTTTTTCGTTTTTTTTT 661
QY 2749 TGAATGTTTTGTTTTTATGTTTTGAGTTCTTGACAGAGATGCTCCCATCATATTTAGT 2808
Db 662 TCGCGTTTTTTTTTTTTCGTTTTTTTTTTTTTTTTTTTTCGAGTTTTTTTTTTTTCGAGT 721
QY 2809 CTTTTTCCTTTCTTTGTTGCTGCTCTCTCTGAGATTTCTCTGATAAAGCTTTACT 2868
Db 722 TTTTTCATTTTTTCGTTTTTTTTTTTTTTTTTTTTCGTTTTTTTTTTTTTTTTCGTTTTT 781
QY 2869 TCTTAACTTTTTCCAGACGGTGAAATTTATTCGATACACCTCAATACCGACCTAT 2928
Db 782 TTTTTCGTTTTGAGTTTTTTTATTTATGAGTTTTTTTTTTTTTTTTCGAGTTTTTTTTTA 841
QY 2929 AATTTAAAGACACTTAGTATGTTTCACTTTTGAAATTTATTTCTATCGAAGGGG 2988
Db 842 ATTTTTTTTTTTTCGGTTTTTTTTTTTCGTTTATCGGTTTTTTTTTTTTTTTTCGAGT 901
QY 2989 CAGATGATTTTCGAGCATGCTCTCGATCATTTTACCGTATTTACAAATGTAATGCGTTGT 3048
Db 902 TTTTTCGTTTTTCGTTTTTTTTTTTTTTTTTTTTCGTTTTTTTTTTTATTTTTTCGTTTT 961
QY 3049 TCGTTTCT 3056
Db 962 TTTTTCGTTTT 969
RESULT 14
US-10-311-455-1669
; Sequence 1669, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR FILING DATE: 2001-07-02
; PRIOR FILING DATE: 2000-06-30
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1669
; LENGTH: 6668
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1099
Query Match 1.5%; Score 59; DB 13; Length 6338;
Best Local Similarity 50.2%; Pred. No. 0.0016;
Matches 146; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
QY 2591 TTGAATTTGTTGAGTCTTTTTTCCCTCATAGGAGTACTATCTACTTCTGATTTTC 2650
Db 4450 TTGGTTTTTTGTTT 4509

ORGANISM: Artificial Sequence
FEATURE: chemically treated genomic DNA (Homo sapiens)
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
NAME/KEY: unsure
LOCATION: 4733
OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1669
Query Match 1.5%; Score 59.6; DB 13; Length 6668;
Best Local Similarity 48.0%; Pred. No. 0.0011;
Matches 170; Conservative 0; Mismatches 184; Indels 0; Gaps 0;
QY 2547 TTATTTACTTACAACTAAGCACCTTTTGTGCTTCGATACACAGTTGAATTTGTTGAGT 2606
Db 5286 TTTTTCGTTTTTGAAGTTTTTGTGTTTTTTTTTTTTTTTTCGTTTCGTTTTTTTTTTT 5345
QY 2607 CTTTTTCCCTCATAGTGGACTAGTCTATTTGTCACATTGATTTTCTTCTCTTTGTTGATGTT 2666
Db 5346 TTTTTCGTTT 5405
QY 2667 TTTCTATGTCATCAAACTCCAAATATGGTAAAGTTTACCTCTTGTGTTGGATTACACAG 2726
Db 5406 TTTTTCGTTTTTATTTGTTGAAGTTTGGTTTTTGGAGTTTTTTTTTTTTTTTAGAAAGT 5465
QY 2727 GTTCCTTTTCATTTCTACACAGTGAATGCTTTTGTGTTTTTATGTTTGTGATTTCTTGACAG 2786
Db 5466 TTTGTTTTTTTTTTTTTTTTCGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 5525
QY 2787 AGATGCTCCCATCATATTTAGTCTCTTTCTCTTTCTCTTCTGTCGTTCTCTCTGATGT 2846
Db 5526 TTTTTCGTTT 5585
QY 2847 TTTCTCTGATAAAGCTTTACTTCTTAACCTTTTTTCCAGCGAGGTTGAATTTAT 2900
Db 5586 TTTTTCGTTTTTTTTTTTTTTTATTTGTTAGATGTTTTTGGAGTTGATTTTTTTTTT 5639
RESULT 15
US-10-311-455-1099
; Sequence 1099, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR FILING DATE: 2001-07-02
; PRIOR FILING DATE: 2000-06-30
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1099
; LENGTH: 6338
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1099
Query Match 1.5%; Score 59; DB 13; Length 6338;
Best Local Similarity 50.2%; Pred. No. 0.0016;
Matches 146; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
QY 2591 TTGAATTTGTTGAGTCTTTTTTCCCTCATAGGAGTACTATCTACTTCTGATTTTC 2650
Db 4450 TTGGTTTTTTGTTT 4509

QY	2651	TTCCCTTTGTTGATGTTTTCTATGTCATGCAAACTCCAAATATGGGTAAAGGTTACCTCCTT	2710
Db	4510		4569
QY	2711	GTTCGGGATTACCAAGATTCCTTTTCATTCTTACACGTGAATGTTGTTTTTATGT	2770
Db	4570		4629
QY	2771	TTTGAGTTCTTGACACAGAGATGCTCCCATCATATTTAGTCCTTTTCCTTTCTCTTTGTGTC	2830
Db	4630		4689
QY	2831	GTTCCTTTCTGGAGTTTCCTTCTGATAAAGCTTTACTTCTTAACTTTTTT	2881
Db	4690		4740

Search completed: January 28, 2004, 21:21:40
Job time : 1246 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 28, 2004, 16:22:39 ; Search time 983 Seconds
(without alignments)
11105.318 Million cell updates/sec

Title: US-10-014-927-18

Perfect score: 4044

Sequence: 1 agacaaagatgcttactct.....atgattatcttaactaac 4044

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_10Jun03.*

- 1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4042.4	100.0	4044	21 AAC81899	A. thaliana SRP30
2	780	19.3	780	24 ABZ16088	Arabidopsis thalia
3	283	7.0	568	21 AAC34694	Arabidopsis thalia
4	281.4	7.0	1213	21 AAC40180	Arabidopsis thalia
5	272.6	6.7	545	25 AEX57250	Arabidopsis thalia
6	222.4	5.5	925	21 AAC49248	Arabidopsis thalia
7	175	4.3	714	21 AAC42865	Arabidopsis thalia
8	143.8	3.6	1041	21 AAC46891	Arabidopsis thalia

c	9	142.2	3.5	919	21	AAC36164	Arabidopsis thalia
	10	117	2.9	289	21	AAC49252	Arabidopsis thalia
	11	117	2.9	771	21	AAC49242	Arabidopsis thalia
	12	108.8	2.7	762	24	ABZ13394	Arabidopsis thalia
c	13	78.6	1.9	5285	19	AAV69915	Arabidopsis ASR-2-
c	14	67.4	1.7	6644	20	AAX33181	Base sequence of t
c	15	67.4	1.7	7372	20	AAX33182	Base sequence of t
c	16	67.4	1.7	7797	20	AAX33180	Cowpox virus bsr f
c	17	67.4	1.7	7996	20	AAX33184	Base sequence of t
18	65.4	1.6	6161	24	ABL32411	Human immune syste	
19	64.8	1.6	14006	24	ABL33958	Human immune syste	
20	64.4	1.6	6171	24	ABL32788	Human immune syste	
21	64.2	1.6	556	24	ABQ36996	Oligonucleotide fo	
c	22	64.2	1.6	556	24	ABQ36997	Oligonucleotide fo
23	63.2	1.6	6668	24	ABL33697	Human immune syste	
c	24	62.4	1.5	875	22	AA195044	Human neuroblastom
25	62.4	1.5	3683	25	AB210199	Haematopoietic cel	
c	26	62	1.5	1257	24	ABZ78031	Human breast speci
27	61.4	1.5	529	25	ABX52180	Bovine EST associa	
28	61.2	1.5	6334	24	ABL33213	Human immune syste	
29	59.8	1.5	885	21	AAC42654	Arabidopsis thalia	
30	59.6	1.5	6668	24	ABL33696	Human immune syste	
31	59.4	1.5	6621	25	ABZ10129	Haematopoietic cel	
32	59.4	1.5	6621	25	ABZ10243	Haematopoietic cel	
33	59.2	1.5	786	24	ABQ33462	Oligonucleotide fo	
c	34	59.2	1.5	786	24	ABQ33463	Oligonucleotide fo
35	59	1.5	6338	24	ABL33126	Human immune syste	
36	58.8	1.5	16633	24	ABN79984	Human immune syste	
37	58.6	1.4	3683	25	AB210053	Human chemically m	
38	58.6	1.4	6767	22	AAS46608	Haematopoietic cel	
39	58.4	1.4	9539	22	AAS45346	Tumour suppressor	
40	58.4	1.4	9539	24	ABK28179	Chemically pretrea	
41	58.2	1.4	8079	24	ABL92313	DNA transcription	
42	58.2	1.4	16033	24	ABL33404	Chemically treated	
43	58.2	1.4	17934	24	ABL33719	Human immune syste	
44	58	1.4	9539	22	AAS45347	Human immune syste	
45	58	1.4	9539	24	ABK28180	Chemically pretrea	
						DNA transcription	

ALIGNMENTS

RESULT 1

AAC81899
ID AAC81899 standard; DNA; 4044 BP.
AC AAC81899;
XX
DT 23-FEB-2001 (first entry)
XX
DE A. thaliana SRP30 cDNA.
XX
KW SR protein; splice-factor activity; plant; developmental behavior;
KW flowering; crop plant; cereal; bean; rice; fruit; ss.
XX
OS Arabidopsis thaliana.
XX
PN WO200065059-A1.
XX
PD 02-NOV-2000.
XX
PF 20-APR-2000; 2000WO-AT00100.
XX
PR 23-APR-1999; 99AT-0000727.
XX
XX (OSTP) OESTERR FORSCH SEIBERSDORF.
XX
XX Barta A, Lopato S, Kalyana M, Dorner S;
XX WPI; 2000-687349/67.
XX P-PSDB; AAB11414, AAB11419.
XX Novel proteins with splice-factor activity in plants, useful e.g. for

Db 1861 TTGTGGATTATAGCAACTATGATGATATGAAAGTACGACGTAAGTATTATATCTTTGCAAC 1920
QY 1921 GCAATGTTCCGGAGCTTAAGCTTAGAGTGGCTTTTGTGTTTCTAGTATATACGAGCTGAAT 1980
Db 1921 GCAATGTTCCGGAGCTTAAGCTTAGAGTGGCTTTTGTGTTTCTAGTATATACGAGCTGAAT 1980
QY 1981 TTATCTTCTCGGAGGCGGAGTGTGGATCTTTGATGTTCCCTTAAATTTTGTGATGAGCA 2040
Db 1981 TTATCTTCTCGGAGGCGGAGTGTGGATCTTTGATGTTCCCTTAAATTTTGTGATGAGCA 2040
QY 2041 GATAAGGAACCTTGATGCCACCTGAATTCGAAATGCTTTCTAGTGGCTTATATACGGGT 2100
Db 2041 GATAGGAACCTTGATGCCACCTGAATTCGAAATGCTTTCTAGTGGCTTATATACGGGT 2100
QY 2101 ATGTTGATGCTTTCTTTGATTTTGTAAAGCATAAAGTGGATGAGTCAFTCTCTGAAT 2160
Db 2101 ATGTTGATGCTTTCTTTGATTTTGTAAAGCATAAAGTGGATGAGTCAFTCTCTGAAT 2160
QY 2161 TTACTGTTTCAGGTGAGGGAATGAGTCCGAGGAGTGTGAGTCCGAGCCCGAGATGTTCTA 2220
Db 2161 TTACTGTTTCAGGTGAGGGAATGAGTCCGAGGAGTGTGAGTCCGAGCCCGAGATGTTCTA 2220
QY 2221 AAAGCTATAGAGCAGGAGTCCGAGGCGGTGTCGAAGCTGTAGCTATAGTAGCAAGAGCA 2280
Db 2221 AAAGCTATAGAGCAGGAGTCCGAGGCGGTGTCGAAGCTGTAGCTATAGTAGCAAGAGCA 2280
QY 2281 GGAGGTCTTTTTTTTTTTTTTTTTTATATAAACCTTAAGACATATAAGGGAATTTTATGTT 2340
Db 2281 GGAGGTCTTTTTTTTTTTTTTTTTTATATAAACCTTAAGACATATAAGGGAATTTTATGTT 2340
QY 2341 AACTTATTTATGAATAAATGACTCTTAATGCAATGCAATGCAAGTGTGTCACTGTAGATCC 2400
Db 2341 AACTTATTTATGAATAAATGACTCTTAATGCAATGCAAGTGTGTCACTGTAGATCC 2400
QY 2401 ATTTCCCGGTTACGGGCCCTTAGTCTGTTCTCGCTCGCTATAGCTATAGCTCTCAAGG 2460
Db 2401 ATTTCCCGGTTACGGGCCCTTAGTCTGTTCTCGCTCGCTATAGCTATAGCTCTCAAGG 2460
QY 2461 TAGAGTGTATGATTTGATTAATTAATATATATATATATATATATATATATATATATAT 2520
Db 2461 TAGAGTGTATGATTTGATTAATTAATATATATATATATATATATATATATATATATAT 2520
QY 2521 TCTTGATAGTGAACCTCTTACTAGCTTTTACTTACTTACTTACTTACTTACTTACTTACTT 2580
Db 2521 TCTTGATAGTGAACCTCTTACTAGCTTTTACTTACTTACTTACTTACTTACTTACTTACTT 2580
QY 2581 CGGTACACAGTTGAATTTGTTGAGTCTTTTCCCTCATATAGTGGAGTCTATTTGTCA 2640
Db 2581 CGGTACACAGTTGAATTTGTTGAGTCTTTTCCCTCATATAGTGGAGTCTATTTGTCA 2640
QY 2641 CTTGATTTTCTTCTTTGTTGATTTTCTATGTCATGCAAACTCCAAATGCGGTAAAGG 2700
Db 2641 CTTGATTTTCTTCTTTGTTGATTTTCTATGTCATGCAAACTCCAAATGCGGTAAAGG 2700
QY 2701 TTACCTCTCTGTTGGGTTTACAGAGTTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2760
Db 2701 TTACCTCTCTGTTGGGTTTACAGAGTTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2760
QY 2761 TTTTATGTTTGTAGTCTTGTACAGAGATGCTCCCATCATATTTAGTCCCTTTCCCTTTC 2820
Db 2761 TTTTATGTTTGTAGTCTTGTACAGAGATGCTCCCATCATATTTAGTCCCTTTCCCTTTC 2820
QY 2821 TCTTTGTGCTGTTCTCTTCTGAGATGTTCTCTGATAAAGCTTTTACTTCTTAACTTTT 2880
Db 2821 TCTTTGTGCTGTTCTCTTCTGAGATGTTCTCTGATAAAGCTTTTACTTCTTAACTTTT 2880
QY 2881 TCCAGCGGCGGGAATTTATAGTATCAACTCACTCAATATCCGACCTTAATTTAAGAAC 2940
Db 2881 TCCAGCGGCGGGAATTTATAGTATCAACTCACTCAATATCCGACCTTAATTTAAGAAC 2940
QY 2941 ACTTACTAGATGTTTCACTTTTGAATTTTATTTCTATCGGAGGCGGCGAGATGTTCT 3000
Db 2941 ACTTACTAGATGTTTCACTTTTGAATTTTATTTCTATCGGAGGCGGCGAGATGTTCT 3000

QY 3001 GAGSCATGTCCTCGATCATTTACCGTATTACAAATTTGATTTGCGTTGTTGTTTCTAGAT 3060
Db 3001 GAGSCATGTCCTCGATCATTTACCGTATTACAAATTTGATTTGCGTTGTTGTTTCTAGAT 3060
QY 3061 CTGGCTCACTGCTACGAGCTGGGATTTGGATCTAGATGGGTCACTAGATGGAATCTTGG 3120
Db 3061 CTGGCTCACTGCTACGAGCTGGGATTTGGATCTAGATGGGTCACTAGATGGAATCTTGG 3120
QY 3121 ACTGGATTTACAAAGCTGGATTTAGATGAACCTGAACTTCTGTTTACGGTCTGGTCTGGT 3180
Db 3121 ACTGGATTTACAAAGCTGGATTTAGATGAACCTGAACTTCTGTTTACGGTCTGGTCTGGT 3180
QY 3181 CTGGTACTCCGCGGTATCAGCTGTAGGATCTGATCGAAAGTTTGGACTATGATTTACT 3240
Db 3181 CTGGTACTCCGCGGTATCAGCTGTAGGATCTGATCGAAAGTTTGGACTATGATTTACT 3240
QY 3241 CTGATTCCTCAATATATTTATCTTTTGAACAATAGTGGATCTGTTGAGTCTCTTTCT 3300
Db 3241 CTGATTCCTCAATATATTTATCTTTTGAACAATAGTGGATCTGTTGAGTCTCTTTCT 3300
QY 3301 AGGACAGCATTTTAAGCTCCCGGACTAGATGGGAGATGTCAGTAAATTTCTTTGTTATG 3360
Db 3301 AGGACAGCATTTTAAGCTCCCGGACTAGATGGGAGATGTCAGTAAATTTCTTTGTTATG 3360
QY 3361 CCACACTTACATGGGGTTTTTGGTCTTGTGTCAGGTCCCAATCAAGATCAAAATCAAGA 3420
Db 3361 CCACACTTACATGGGGTTTTTGGTCTTGTGTCAGGTCCCAATCAAGATCAAAATCAAGA 3420
QY 3421 TCAAGATCAAGATCGAATCTCCAGTTTCCAGTTTCCAGTTTCCAGTTTCCAGTTTCCAGTTT 3480
Db 3421 TCAAGATCAAGATCGAATCTCCAGTTTCCAGTTTCCAGTTTCCAGTTTCCAGTTTCCAGTTT 3480
QY 3481 TTAATTCACAAATCATGTTGTTTAAATACCTGCTCACTTTGGTTGTTTCTTCAATCAA 3540
Db 3481 TTAATTCACAAATCATGTTGTTTAAATACCTGCTCACTTTGGTTGTTTCTTCAATCAA 3540
QY 3541 CACCAACTTAACGAAATPCATGAGACAGACTATAAAATTTGAAGAGTCTGTAGAACGACTA 3600
Db 3541 CACCAACTTAACGAAATPCATGAGACAGACTATAAAATTTGAAGAGTCTGTAGAACGACTA 3600
QY 3601 GGTCTCACCAACTCTGTGTGCACTAAATCGCTCTCCAGTGTTCAGCAACATAAT 3660
Db 3601 GGTCTCACCAACTCTGTGTGCACTAAATCGCTCTCCAGTGTTCAGCAACATAAT 3660
QY 3661 CTACCTCTGTCATGTTATCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3720
Db 3661 CTACCTCTGTCATGTTATCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3720
QY 3721 CAGGTGATATCTGTTTGAATGAAACTGGCCACTGGCTGTACCCGAATCGTCTCAAGC 3780
Db 3721 CAGGTGATATCTGTTTGAATGAAACTGGCCACTGGCTGTACCCGAATCGTCTCAAGC 3780
QY 3781 TTTCTCAGGCTCCACTGCTAATAGAAATTTGATTCGATTTGGGATTTATATATGTTCTTC 3840
Db 3781 TTTCTCAGGCTCCACTGCTAATAGAAATTTGATTCGATTTGGGATTTATATATGTTCTTC 3840
QY 3841 TTGTATGGGAGCAATATGCTTTCTAGTCTTTTGTAGTGTGAACCTGGAAATGGTCTGTT 3900
Db 3841 TTGTATGGGAGCAATATGCTTTCTAGTCTTTTGTAGTGTGAACCTGGAAATGGTCTGTT 3900
QY 3901 ATTGTGTCATTAAAGCCGAAACTCTGCTCGGCTGCAATAATAAGTTTCAATCAGACAT 3960
Db 3901 ATTGTGTCATTAAAGCCGAAACTCTGCTCGGCTGCAATAATAAGTTTCAATCAGACAT 3960
QY 3961 TGTGTTGGGTGCTGAGGTTTTTCCATACATATATATATATATATATATATATATATATAT 4020
Db 3961 TGTGTTGGGTGCTGAGGTTTTTCCATACATATATATATATATATATATATATATATATAT 4020
QY 4021 TTTTATGATTTCTTAAACTAAAC 4044
Db 4021 TTTTATGATTTCTTAAACTAAAC 4044

PR 23-APR-1999;	99US-0130510.
PR 23-APR-1999;	99US-0130891.
PR 28-APR-1999;	99US-0134449.
PR 30-APR-1999;	99US-0132048.
PR 30-APR-1999;	99US-0132407.
PR 04-MAY-1999;	99US-0132484.
PR 05-MAY-1999;	99US-0132485.
PR 06-MAY-1999;	99US-0132486.
PR 06-MAY-1999;	99US-0132487.
PR 07-MAY-1999;	99US-0132863.
PR 11-MAY-1999;	99US-0134256.
PR 14-MAY-1999;	99US-0134218.
PR 14-MAY-1999;	99US-0134219.
PR 14-MAY-1999;	99US-0134221.
PR 14-MAY-1999;	99US-0134370.
PR 18-MAY-1999;	99US-0134768.
PR 19-MAY-1999;	99US-0134941.
PR 20-MAY-1999;	99US-0135124.
PR 21-MAY-1999;	99US-0135353.
PR 24-MAY-1999;	99US-0135629.
PR 25-MAY-1999;	99US-0136021.
PR 27-MAY-1999;	99US-0136392.
PR 28-MAY-1999;	99US-0136782.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137724.
PR 08-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 10-JUN-1999;	99US-0138847.
PR 14-JUN-1999;	99US-0139119.
PR 16-JUN-1999;	99US-0139452.
PR 16-JUN-1999;	99US-0139453.
PR 17-JUN-1999;	99US-0139492.
PR 18-JUN-1999;	99US-0139454.
PR 18-JUN-1999;	99US-0139455.
PR 18-JUN-1999;	99US-0139456.
PR 18-JUN-1999;	99US-0139457.
PR 18-JUN-1999;	99US-0139458.
PR 18-JUN-1999;	99US-0139459.
PR 18-JUN-1999;	99US-0139460.
PR 18-JUN-1999;	99US-0139461.
PR 18-JUN-1999;	99US-0139462.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 18-JUN-1999;	99US-0139763.
PR 21-JUN-1999;	99US-0139817.
PR 22-JUN-1999;	99US-0139899.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140695.
PR 28-JUN-1999;	99US-0140823.
PR 29-JUN-1999;	99US-0140991.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141842.
PR 02-JUL-1999;	99US-0142154.
PR 06-JUL-1999;	99US-0142055.
PR 06-JUL-1999;	99US-0142390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0142977.
PR 13-JUL-1999;	99US-0143342.
PR 14-JUL-1999;	99US-0143624.
PR 15-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144085.
PR 18-JUL-1999;	99US-0144086.
PR 18-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144332.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 19-JUL-1999;	99US-0144335.
PR 20-JUL-1999;	99US-0144352.
PR 20-JUL-1999;	99US-0144632.
PR 20-JUL-1999;	99US-0144884.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145087.
PR 22-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 26-JUL-1999;	99US-0145224.
PR 27-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 27-JUL-1999;	99US-0145919.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 06-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 09-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149358.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159233.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.

T 4
180
AAC40180 standard; DNA; 1213 BP.
AAC40180;
17-OCT-2000 (first entry)
Arabidopsis thaliana DNA fragment SEQ ID NO: 27352.
Hybridisation assay; genetic mapping; gene expression control;
protein identification; signal transduction pathway;
metabolic pathway; promoter; termination sequence; ss.
Arabidopsis thaliana.
EP1033405-A2.
06-SEP-2000.
25-FEB-2000; 2000EP-0301439.
25-FEB-1999; 99US-0121825.
05-MAR-1999; 99US-0123180.
09-MAR-1999; 99US-0123548.

[illegible]

PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144684.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.

PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149829.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	22-SEP-1999;	99US-0154779.
PR	23-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159638.
PR	21-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	22-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161320.
PR	28-OCT-1999;	99US-0161392.
PR	28-OCT-1999;	99US-0161393.
PR	29-OCT-1999;	99US-0162142.

Query Match

5.5%; Score 222.4; DB 21; Length 925;

[illegible]

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140635.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0148171.
PR 10-AUG-1999; 99US-0148319.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

```

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.6%; Score 143.8; DB 21; Length 1041;
Best Local Similarity 98.6%; Pred. No. 1.6e-24;
Matches 145; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 229 CTCACGACTGAGAGACCGTGCGTGTCTACGGCGGAAATGCTCTTGACCGAGTT 288
Db 150 CTTCCGACTGAGAGACCGTGCGTGTCTACGGCGGAAATGCTCTTGACCGAGTT 91
QY 289 ACTAATCGGCAATTGAGATTTGAGAGGTTGAGAGGAGAAATGTTTCTCT 348
Db 90 ACTAATCGGCAATTGAGATTTGAGAGGTTGAGAGGAGAAATGTTTCTCT 31

QY 349 CAAAAAAATCCCAAGTGTTCGATC 375
Db 30 CAAAAAAATCCCAAGTGTTCGATC 4

RESULT 9
AAC36164/c
ID AAC36164 standard; DNA; 919 BP.
XX
AC AAC36164;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 12774.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN BP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0140931.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144854.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.

```

PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151203.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0159369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.5%; Score 142.2; DB 21; Length 919;
Best Local Similarity 98.0%; Pred. No. 3.8e-24;
Matches 144; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 229 CTCACGGACTGAGAGACCGTGGACGGTGTCAACGGCGGAAAATGCTCTTGGACCGAGTT 288
|||
Db 147 CTTCCGACTGAGAGACCGTGGACGGTGTCAACGGCGGAGAATGCTCTTGGACCGAGTT 88
|||

Qy 289 ACTAATCGGCGAATTCAGATTTCAGAGGTGGTAGAGGACCGGACCGAGAGAATGTTTCTCT 348
|||
Db 87 ACTAATCGGCGAATTCAGATTTCAGAGGTGGTAGAGGACCGGACCGAGAGAATGTTTCTCT 28
|||

Qy 349 CAAAAAATCCCCAAGTGTTCGGATC 375
|||
Db 27 CAAAAAATCCCCAAGTGTTCGGATC 1
|||

RESULT 10
AAC49252
ID AAC49252 standard; DNA; 289 BP.
XX AAC49252;
AC
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 60485.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
CS
XX
PN BP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 05-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.

PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148311.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139456.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139457.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139463.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	13-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0142154.	PR	16-SEP-1999;	99US-0154039.
PR	02-JUL-1999;	99US-0142055.	PR	20-SEP-1999;	99US-0154339.
PR	06-JUL-1999;	99US-0142390.	PR	22-SEP-1999;	99US-0155139.
PR	08-JUL-1999;	99US-0142803.	PR	23-SEP-1999;	99US-0155139.
PR	09-JUL-1999;	99US-0142920.	PR	23-SEP-1999;	99US-0155479.
PR	12-JUL-1999;	99US-0142977.	PR	28-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	29-SEP-1999;	99US-0156596.
PR	14-JUL-1999;	99US-0143624.	PR	04-OCT-1999;	99US-0157117.
PR	15-JUL-1999;	99US-0144005.	PR	05-OCT-1999;	99US-0157753.
PR	16-JUL-1999;	99US-0144085.	PR	06-OCT-1999;	99US-01

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.9%; Score 117; DB 21; Length 289;
Best Local Similarity 89.4%; Pred. No. 2.9e-18;
Matches 126; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 2152 TCTCTGAATTTACTGTTTCAGGTGAGGGAATATGATGTCGAGSAGTGTGAGTCGAAGCCAG 2211
Db 18 TCTCTAGTCTTATATACGGGTGAGGGAATATGATGTCGAGSAGTGTGAGTCGAAGCCAG 77

Qy 2212 ATGATTTCTAAAGCTATAGAAGCAGGAGTCGGAGCCGTGTCCAAAGCTGTAGCTATAGTA 2271
Db 78 ATGATTTCTAAAGCTATAGAAGCAGGAGTCGGAGCCGTGTCCAAAGCTGTAGCTATAGTA 137

Qy 2272 GCAAGAGCAGGAGCTCTTTT 2292
Db 138 GCAAGAGCAGGAGCTGTGT 158

RESULT 11
AAC49242
ID AAC49242 standard; DNA; 771 BP.
AC AAC49242;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 60449.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127482.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-013817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

23-JUL-1999; 99US-0145224.
 26-JUL-1999; 99US-0145276.
 27-JUL-1999; 99US-0145313.
 27-JUL-1999; 99US-0145318.
 27-JUL-1999; 99US-0145319.
 28-JUL-1999; 99US-0145351.
 02-AUG-1999; 99US-0145386.
 02-AUG-1999; 99US-0146388.
 02-AUG-1999; 99US-0146389.
 03-AUG-1999; 99US-0147038.
 04-AUG-1999; 99US-0147204.
 04-AUG-1999; 99US-0147302.
 05-AUG-1999; 99US-0147192.
 05-AUG-1999; 99US-0147260.
 06-AUG-1999; 99US-0147303.
 06-AUG-1999; 99US-0147416.
 08-AUG-1999; 99US-0147493.
 08-AUG-1999; 99US-0147935.
 10-AUG-1999; 99US-0148171.
 11-AUG-1999; 99US-0148319.
 12-AUG-1999; 99US-0148341.
 13-AUG-1999; 99US-0148565.
 13-AUG-1999; 99US-0148684.
 16-AUG-1999; 99US-0149368.
 17-AUG-1999; 99US-0149175.
 17-AUG-1999; 99US-0149178.
 18-AUG-1999; 99US-0149426.
 20-AUG-1999; 99US-0149722.
 20-AUG-1999; 99US-0149723.
 20-AUG-1999; 99US-0149929.
 23-AUG-1999; 99US-0149902.
 23-AUG-1999; 99US-0149930.
 25-AUG-1999; 99US-0150566.
 26-AUG-1999; 99US-0150884.
 26-AUG-1999; 99US-0151065.
 27-AUG-1999; 99US-0151066.
 27-AUG-1999; 99US-0151080.
 30-AUG-1999; 99US-0151303.
 31-AUG-1999; 99US-0151438.
 01-SEP-1999; 99US-0151930.
 07-SEP-1999; 99US-0152363.
 10-SEP-1999; 99US-0153070.
 13-SEP-1999; 99US-0153758.
 15-SEP-1999; 99US-0154018.
 16-SEP-1999; 99US-0154039.
 20-SEP-1999; 99US-0154779.
 22-SEP-1999; 99US-0155139.
 23-SEP-1999; 99US-0155486.
 24-SEP-1999; 99US-0155659.
 28-SEP-1999; 99US-0156458.
 29-SEP-1999; 99US-0156596.
 04-OCT-1999; 99US-0157117.
 05-OCT-1999; 99US-0157753.
 06-OCT-1999; 99US-0157865.
 07-OCT-1999; 99US-0158029.
 08-OCT-1999; 99US-0158232.
 12-OCT-1999; 99US-0158369.
 13-OCT-1999; 99US-0158293.
 13-OCT-1999; 99US-0159294.
 13-OCT-1999; 99US-0159295.
 14-OCT-1999; 99US-0159329.
 14-OCT-1999; 99US-0159330.
 14-OCT-1999; 99US-0159331.
 14-OCT-1999; 99US-0159637.
 14-OCT-1999; 99US-0159638.
 18-OCT-1999; 99US-0159584.
 21-OCT-1999; 99US-0160741.
 21-OCT-1999; 99US-0160767.
 21-OCT-1999; 99US-0160768.
 21-OCT-1999; 99US-0160770.
 21-OCT-1999; 99US-0160814.
 21-OCT-1999; 99US-0160815.
 22-OCT-1999; 99US-0160980.
 22-OCT-1999; 99US-0160981.

22-OCT-1999; 99US-0160989.
 25-OCT-1999; 99US-0161404.
 25-OCT-1999; 99US-0161405.
 25-OCT-1999; 99US-0161406.
 26-OCT-1999; 99US-0161359.
 26-OCT-1999; 99US-0161360.
 26-OCT-1999; 99US-0161361.
 28-OCT-1999; 99US-0161920.
 28-OCT-1999; 99US-0161992.
 28-OCT-1999; 99US-0161993.
 29-OCT-1999; 99US-0162142.

Query-Match 2.9%; Score 117; DB 21; Length 771;
 Best Local Similarity 92.5%; Pred. No. 4.3e-18;
 Matches 123; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2152 TCCTGAAATTTACTGTTTCAGCTGAGGGAATATGATCGAGGAGTGTGAGTCGAAGCCCGAG 2211
 |||||
 DB 524 TCCTGAGTCTTATATACGGTGGAGGAATATGATCGAGGAGTGTGAGTCGAAGCCCGAG 583
 |||||
 QY 2212 ATGATTCTAAAAGCTATAGAAGCAGGAGTCGAGCCCGTGGTCCAAAGCTGTAGCTATAGTA 2271
 |||||
 DB 584 ATGATTCTAAAAGCTATAGAAGCAGGAGTCGAGCCCGTGGTCCAAAGCTGTAGCTATAGTA 643
 |||||

QY 2272 GCAGAGCAGGAG 2284
 |||||
 DB 644 GCAGAGCAGGAG 656
 |||||

RESULT 12
 ABZ13394
 ID ABZ13394 standard; DNA; 762 BP.
 XX AC ABZ13394;
 AC ABZ13394;
 XX DT 21-JAN-2003 (first entry)
 XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1199.
 XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
 XX OS Arabidopsis thaliana.
 XX PN WO200216655-A2.
 XX PD 28-FEB-2002.
 XX PF 24-AUG-2001; 2001WO-US26685.
 XX PR 24-AUG-2000; 2000US-227866P.
 XX PR 26-JAN-2001; 2001US-264647P.
 XX PR 22-JUN-2001; 2001US-300111P.
 XX (SCRI) SCRIPPS RES INST.
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX PI Harper JF, Kreps J, Wang X, Zhu T;
 XX WPI; 2002-304127/34.
 DR Identifying a stress condition to which a plant cell has been exposed
 XX and producing plants with increased tolerance to these abiotic stresses
 PT -
 XX Claim 144; SEQ ID NO 1199; 577pp + Sequence Listing; English.
 PS The invention relates to identifying a stress condition to which a plant
 XX cell has been exposed, comprising:
 CC (a) contacting nucleic acid representative of expressed polynucleotides
 CC in the plant cell with an array or probes representative of the plant
 CC cell genome; and
 CC (b) detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the plant cell

CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (AB212198-AB211574) used
CC in methods of the invention.

CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.

Sequence 762 BP; 189 A; 161 C; 197 G; 215 T; 0 other;

	Query Match	2.7%; Score 108.8; DB 24; Length 762;
	Best Local Similarity 90.6%;	Pred. No. 4.1e-16;
	Matches 116; Conservative 0;	Mismatches 12; Indels 0; Gaps 0;
Qy	1300	TGTTATCTTTGTGTGATGTTTTTATAGTTTGAGATTGCACATGGTGGTCGTAGATTTCACC 1358
Db	204	TGATTTTGAAGGTTGCGACTTCGGTTGAGATTGCATGFTGGTCGTAGATTTCACC 263
Qy	1360	ATCAGTTGATAGGTACAGCAGCAGCTACGTGGGAGCGGTGCACCTTCAAGACGCTCTGA 1419
Db	264	ATCAGTTGATAGGTACAGCAGCAGCTACGTGGGAGCGGTGCACCTTCAAGACGCTCTGA 323
Qy	1420	CTACCGCG 1427
Db	324	CTACCGCG 331

RESULT 13
AAV69915/C
ID AAV69915 standard: DNA: 5285 BP.

DT 01-MAR-1999 (first entry)

DE Arabidopsis ASR-2-ORF3 gene region.

Promoter: transgenic plant: ASR-2 gene: vector: ds.

OS *Arabidopsis thaliana* ecotype RDL.

PH	Key	Location/Qualifiers
FT	CDS	945..3694
FT		/*tag= a
FT		/label= ASR-2
FT	promoter	complement (3691..4220)
FT		/*tag= b
FT		/note= "Claim 1"
FT	CDS	complement (4217..4917)
FT		/*tag= c
FT		/label= CRE3

PN WO9844781-A1.

15-OCT-1998.

03-APR-1998;

PR 04-APR-1997; 97US-0042926.

PA (PURD) PURDUE RES FOUND.

PI Hodges TK, Lyznik LA;

DR WPI; 1998-609898/51.

PT Novel promoter sequences for

PT sequence useful for high exp

PS Example 1; Page 11-14; 23pp;

CC This is the nucleotide sequence

of *Arabidopsis thaliana*. A genomic library of *A. thaliana* ecotype RDL was screened with a BglI fragment of the rice anther-specific cDNA clone KIS-1. A genomic fragment was isolated and identified as containing sequences homologous to the human splice factor ASR/SP2 and the *Arabidopsis* SRI gene, and was designated ASR-2. The ASR-2 gene was found to be expressed in all plant parts investigated. The genomic fragment also included ORF3 on the opposite strand relative to ASR-2. A 530 bp region located between the 2 genes was shown to function as a dual promoter. The claimed promoter (see AAV65913) is in the orientation that naturally expresses the ORF3 gene. It is capable of driving strong expression of linked exogenous genes. The claimed promoter can be incorporated into expression vectors to drive high level protein expression in transgenic plants. The exogenous gene may encode products that provide e.g. herbicide, insect or fungal resistance, or which act as growth regulators, or which encode pharmaceutical or polymer components. The promoter is equally, or more, effective than the 35S cauliflower mosaic virus promoter.

Sequence 5285 BP; 1702 A; 1063 C; 1066 G; 1454 T; 0 other;

Query Match	1.9%;	Score 78.6;	DB 19;	Length 5285;	
Best Local Similarity	51.1%;	Pred. No. 1.7e-08;			
Matches 408;	Conservative	0;	Mismatches 329;	Indels 62;	Gaps 7;
Qy	1159	TATTGGTGGCAGTTTGAAGATCTCTGCTGATGCACAGATCAATTTATGACCGTATGGT	1218		
Db	3225	TATGTGATTAGTTTGGAGATGCTCGTGATGCTGATGATGCANTTTATGSCCGTATGGT	3166		
Qy	1219	TATGATTTTGATCGGTGTGACATTCGGTTPAGTAAA-----CGCATGATGAAAG	1267		
Db	3165	TATGACTTTGATGGGCATCATTTACGGGTTGTAAATCATCAAAATTTCAGAATTTTACA	3106		
Qy	1268	CTAGCTTAATTTTCTGTAATTTCTTGTAAGGTTTATCTTCTGG-TGATGTTTTTAGGT	1326		
Db	3105	CAAACTTTTGATTGTTTGACTGTPAGCCAAATAATTTATGTGTTTGGCCATTGCAGGT	3046		
Qy	1327	TGAGATTGACATGCTGGTGGTTCGTAGATTTTCCAC---CATCAGTTTGATAGGTACAGCAGCAG	1383		
Db	3045	GGAACTAGCTCATGTTGGGAGGGCTCATCACATGATGCACGGTAGTTTATAGTGTCTG	2986		
Qy	1384	CTACAGTGCAGCCGTGC-----ACCTTCAAGACGCCTCTGAC	1420		
Db	2985	TGGTCTGGTCGGTTCGTTGGTGTGACGGTGTGGTGCCTGAAACGTGGACCATCTAGGAG	2926		
Qy	1421	TACGGCGTTTGAGAGTCTTCTCGATTGTTATTTGGTGTGTGTGTAAATTTTATATT	1480		
Db	2925	ATCAGAGTACCGCGGTACATATGATATGTTTAGCTTTAGTGTATTCATAAGTTTTAGGAG	2866		
Qy	1481	TGAAACTCATTTTTTACTACCTAAACATGATGATGCTGTGCACCGGATTACCGCCTCTGCG	1540		
Db	2865	AAATCATCTAAATTAAGATNTTCCACTATACAGTTGTAGTGTACAGTTTGGCTTCATCTGC	2806		
Qy	1541	TTCGTGGCAGGACCTTAAAGTAAAGGCACATATATAGTCTTTTTCTCTGAATGTGTGTTTC	1600		
Db	2805	GTCTGGCAGAGCCTCAAGGTGATTACACAGTTGTGGACAGATTC-----CTCATTTCTCG	2751		
Qy	1601	TCTATATCATGTTTTGGATTTATCTCTTTTCTGAAATGATGTTATTTGCTATTATTCGGGT	1660		
Db	2750	TTCTATATTTTCTACGATTTCTGAATGTTTGACATATCACTTTTGTTT-----	2704		
Qy	1661	GATTAGGATCACATGCGCAAGCTGGAGATGTCTGCTTCTCTGAAGTTTTCCCTGACCGGT	1720		
Db	2703	TGAGGATCACATGCGCTAAAGGAGGAAGTTTGTTTTTTCTCAAGTGTTCGTGATGGT	2645		
Qy	1721	AAAGGTGAGTTGCATTCGATAGTTTGGTAAGCTTTTTTGATTCGATGCTGTAGTAAATTA	1780		
Db	2644	AGAGTAAATCCCTCATGCTCTTACCCAGAATAGAGATGTCTCTAAAGCTGTTTGGGTTTT	2585		
Qy	1781	GTCTTTGTGAAGAGAAATAGGTTTAAGCATCTGAACCTGCTAAA-----CTCACATTCAG	1835		
Db	2584	CTGCTTTAAATCCTGTTTTTAACTAGTCTCTGATGCTTTTAAAAAAATCTATATCTGGCAG	2525		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 28, 2004, 17:00:12 ; Search time 7850 Seconds

(without alignments)
12520.675 Million cell updates/sec

Title: US-10-014-927-18

Perfect score: 4044

Sequence: 1 agacaaagatcttacttct.....atgattatcttaactaaac 4044

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

BST:

1: em_estba:**

2: em_esthum:**

3: em_estin:**

4: em_estmu:**

5: em_estov:**

6: em_estpl:**

7: em_estro:**

8: em_hic:**

9: gb_est1:**

10: gb_est2:**

11: gb_hic:**

12: gb_est3:**

13: gb_est4:**

14: gb_est5:**

15: em_estfun:**

16: em_estom:**

17: em_ges_hum:**

18: em_ges_inv:**

19: em_ges_pln:**

20: em_ges_vrt:**

21: em_ges_fun:**

22: em_ges_mam:**

23: em_ges_mus:**

24: em_ges_pro:**

25: em_ges_rod:**

26: em_ges_phg:**

27: em_ges_vrl:**

28: gb_ges1:**

29: gb_ges2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	571.4	14.1	1020	28	Bi0994 F7G19-Sp6.3
C 2	529	13.1	867	28	B08242 F23J11-Sp6
C 3	320.4	7.9	928	28	Bi0052 F23J11-Sp6
4	309	7.6	973	28	Bi1569 F3E19-Sp6 I

C 5	305	7.5	445	9	AV794356
C 6	272.4	6.7	510	9	AV531252
C 7	265	6.6	276	29	BZ358568
8	263.4	6.5	1044	28	Bi11684
9	263	6.5	925	28	Bi12280
10	260.2	6.4	340	29	BZ592759
11	247.8	6.1	699	28	Bi12780
C 12	246.4	6.1	791	29	BZ471603
C 13	243.4	6.0	298	29	AL950061
C 14	235.6	5.8	300	9	AV532256
C 15	227.4	5.6	439	9	AI998038
C 16	223.8	5.5	694	28	BZ065015
C 17	209.8	5.2	631	28	BZ065073
C 18	207.2	5.1	467	9	AM004187
C 19	207.2	5.1	676	29	BZ433601
C 20	206.6	5.1	293	29	AL950047
C 21	203.4	5.0	277	14	T42588
C 22	202.4	5.0	411	14	R65514
C 23	188	4.6	199	29	BZ381462
C 24	187	4.6	762	28	BZ062593
C 25	178.8	4.4	741	28	BH589916
C 26	165.4	4.1	186	29	BZ381461
C 27	165.4	4.1	499	9	AI994618
C 28	132.8	3.3	311	14	Z18005
C 29	117	2.9	604	9	AV826310
C 30	113.8	2.8	675	9	AV821889
C 31	109.2	2.7	149	29	AL757304
C 32	98.6	2.4	361	28	BH757020
C 33	97.8	2.4	571	28	AQ962088
C 34	96.4	2.4	567	28	AQ962089
C 35	95.2	2.4	758	28	BH540746
C 36	90.8	2.2	762	28	BH559250
C 37	90.4	2.2	826	29	BZ471593
C 38	89.2	2.2	126	29	AL757303
C 39	86.6	2.1	724	28	BH479317
C 40	81.4	2.0	1078	29	CNS014EL
C 41	80.6	2.0	997	14	CD247514
C 42	79.8	2.0	741	29	CNS007YN
C 43	78	1.9	1132	29	CC248307
C 44	77.6	1.9	975	13	BU960792
C 45	76.6	1.9	1043	14	CD385012

ALIGNMENTS

RESULT 1
Bi0994
LOCUS F7G19-Sp6.3 IGF Arabidopsis thaliana genomic clone F7G19, genomic survey sequence.
DEFINITION F7G19-Sp6.3 IGF Arabidopsis thaliana genomic clone F7G19, genomic survey sequence.
ACCESSION Bi0994
VERSION Bi0994.1 GI:2092116
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; euroids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 1020)
AUTHORS Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and Ecker, J.
TITLE BAC End Sequences at ATGC
JOURNAL Unpublished
COMMENT Other GSSs: F7G19-T7.2, F7G19-T7.1, F7G19-T7.4, F7G19-T7.3, F7G19-Sp6.2, F7G19-T7, F7G19-Sp6, F7G19-Sp6.1
Contact: Ecker, J.
Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104
Tel: 215-898-9384
Fax: 215-898-8780

Email: jecker@atgenome.bio.upenn.edu
Seq primer: Sp6
Class: BAC ends
High quality sequence start: 81
High quality sequence stop: 969.
Location/Qualifiers

FEATURES

Source
1..1020
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="F7G19"
/sex="hermaphrodite"
/clone_lib="IGF"
/note="Vector: BelobACII; Site_1: EcoRI; Site_2: EcoRI;
Produced by Thomas Altmann"

BASE COUNT 283 a 234 c 166 g 331 t 6 others

Query Match 14.1%; Score 571.4; DB 28; Length 1020;
Best Local Similarity 99.0%; Pred. No. 2.4e-73;
Matches 607; Conservative 0; Mismatches 1; Indels 5; Gaps 3;

QY 3433 TCGAATTCACGATTCACCTGGTGAAGTCTAAAGCTGAACCTCTTTAAATCAAT 3492
Db 69 TCGAATTCACGATTCACCTGGTGAAGTCTAAAGCTGAACCTCTTTAAATCAAT 125
QY 3493 CCAT-GTGTGTTTAAATACCTGCTCACTTGGTGTCTTCAATCAACCAACTAA 3551
Db 126 CCATGGTGTGTTTAAATACCTGCTCACTTGGTGTCTTCAATCAACCAACTAA 185
QY 3552 CGAATCATGAGACAGACTATAAATTTGAAGAGTCTGTAGAAGCAGTAGTCTCAACAA 3611
Db 186 CGAATCATGAGACAGACTATAAATTTGAAGAGTCTGTAGAAGCAGTAGTCTCAACAA 244
QY 3612 CCTGTGTGCACTAAATCGCTCTCCAAAGTCTTCAAGCAATATCTACCTGTGTC 3671
Db 245 CCTGTGTGCACTAAATCGCTCTCCAAAGTCTTCAAGCAATATCTACCTGTGTC 304
QY 3672 ATGTGTTATCATTTCTTCTTCTTAACGGTATTACATATATATGTTTGCAGGTGATC 3731
Db 305 ATGTGTTATCATTTCTTCTTCTTAACGGTATTACATATATATGTTTGCAGGTGATC 364
QY 3732 TGGTGAATATGAATGAGCTGGCTGCTGACCGAATGCTCTCAAGCTTCTCAGGCTC 3791
Db 365 TGGTGAATATGAATGAGCTGGCTGCTGACCGAATGCTCTCAAGCTTCTCAGGCTC 424
QY 3792 CACTGCTAATAGAAATTTGATTCGGATTGGGATTATATATGTTTGTATGGGAC 3851
Db 425 CACTGCTAATAGAAATTTGATTCGGATTGGGATTATATATGTTTGTATGGGAC 484
QY 3852 GACCAATATGCTTTCTAGTTTATGTTGAACTGGAATGGTCTGTTATGTCATT 3911
Db 485 GACCAATATGCTTTCTAGTTTATGTTGAACTGGAATGGTCTGTTATGTCATT 544
QY 3912 AAAAGCGGAACTGCTCGCTGCATATTAAGTCTATCAGCAATGTTGGTG 3971
Db 545 AAAAGCGGAACTGCTCGCTGCATATTAAGTCTATCAGCAATGTTGGTG 604
QY 3972 TGGTGAGTGTTCATACATATACATTTACATTAACACTGCTGCTTTTATGATTA 4031
Db 605 TGGTGAGTGTTCATACATATACATTTACATTAACACTGCTGCTTTTATGATTA 664
QY 4032 TCTTAAACTAAAC 4044
Db 665 TCTTAAACTAAAC 677

RESULT 2
B08242/c
LOCUS F23J11-Sp6.1 IGF Arabidopsis thaliana genomic clone F23J11, genomic
DEFINITION survey sequence.
867 bp DNA linear GSS 14-MAY-1997

ACCESSION

B08242
B08242.1 GI:2089364

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

AUTHORS

Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and
Ecker, J.

TITLE

JOURNAL

COMMENT

BAC End Sequences at ATGC
Unpublished
Other GSSs: F23J11-Sp6, F23J11-T7.1, F23J11-T7
Contact: Ecker J.
Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA
19104

Tel: 215-898-9384

Fax: 215-898-8780

Email: jecker@atgenome.bio.upenn.edu

Seq primer: Sp6

Class: BAC ends

High quality sequence start: 48

High quality sequence stop: 549.

FEATURES

source

1..867
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="F23J11"
/sex="hermaphrodite"
/clone_lib="IGF"
/note="Vector: BelobACII; Site_1: EcoRI; Site_2: EcoRI;
Produced by Thomas Altmann"

BASE COUNT 290 a 207 c 153 g 209 t 8 others

Query Match 13.1%; Score 529; DB 28; Length 867;
Best Local Similarity 82.8%; Pred. No. 3.6e-67;
Matches 684; Conservative 0; Mismatches 132; Indels 10; Gaps 7;

QY 2647 TTCTTCCTTTGTCATGTTTCTATGTCATGCAACCTCCATATGGGTAAAGGTACT 2706
Db 838 TTCTTCCTTTGTCATGTTTCTATGTCATGCAACCTCCATATGGGTAAAGGTACT 779
QY 2707 CCTTGTGTTGGGATTACAGAGTTCCTTTTCTTACAGTGAATGTTGTTT 2766
Db 778 CCTTGTGTTGGGATTACAGAGTTCCTTTTCTTACAGTGAATGTTGTTT 719
QY 2767 ATGTTTTCAGTTCCTTGACAGAGATGCTCCCATCATATTTAGTCTTCTTCTTTG 2826
Db 718 AGTTTTCAGTTCCTTGACAGAGATGCTCCCATCATATTTAGTCTTCTTCTTTG 659
QY 2827 TGTGTTCTTCTGATGTTTCTTCTGATGTTTCTTCTGATGTTTCTTCTTCTTCTT 2886
Db 658 GGTTCATCTCTTNGGAGGTTCTTCTGATGTTTCTTCTGATGTTTCTTCTTCTTCTT 599
QY 2887 GACGTTGATTTATACGATCAACCTCAATCGACCTATATTTAAAGACACTTAG 2946
Db 598 GACGTTGATTTATACGATCAACCTCAATCGACCTATATTTAAAGACACTTAG 540
QY 2947 CTAGATGTT---CACTTTTGAATAATTTATTTCTATCGGAAGGGGCA--GATGATTTCTG 3001
Db 539 GTAGATGTTTCACTTTTGAATAATTTATTTCTATCGGAAGGGGCA--GATGATTTCTG 480
QY 3002 AGGCATGTCCTCGATCATTTACCGTGATTAACAATGATTCGTTGTTG--TTTCTAGAT 3060
Db 479 AGGCATGTCCTCGATCATTTACCGTGATTAACAATGATTCGTTGTTG--TTTCTAGAT 420
QY 3061 CTGGCTCACTGCTAGCAGCTGGGATTTGATCTAGATGGGTCACTAGATGTTCTTGG 3120

Db 419 CTGGCTCACTGCTACGAGCTGGGATTGGATCTAGATGGTCACTAGATGGATCTTGG 360
QY 3121 ACTGATTTACAAAGCTGATAGCAAC-TGAACCTCTG-TTTTACGGTCTGGCTG 3178
Db 359 ACTGATTTACAAAGCTGATAGCAAC-TGAACCTCTG-TTTTACGGTCTGGCTG 300
QY 3179 GCTGATCTCCCGGGTATCAGCTGTAGATCTGATCGCAAGTTTGGACTATGATTA 3238
Db 299 GTCGTGATCTCCCGGGTATCAGCTGTAGATCTGATCGCAAGTTTGGACTATGATTA 240
QY 3239 CTCGTATCTCAATATATTTATCTTTTGACAAAGTGTGTTGTTAGTTCTTTT 3298
Db 239 CTCGTATCTCAATATATTTATCTTTTGACAAAGTGTGTTGTTAGTTCTTTT 180
QY 3299 CTAGACAGCATTTAAGCTCCCGGACTAGATGGGATGGTCACTAAATTTCTTTGTTA 3358
Db 179 CTAGACAGCATTTAAGCTCCCGGACTAGATGGGATGGTCACTAAATTTCTTTGTTA 120
QY 3359 TGGCAGACTATAGGGGTTTTTGGTCTTGTGTCAGGTCCTCAATCAAGATCAAAATCAA 3418
Db 119 TGGCAGACTATAGGGGTTTTT-GGTCTGTGTCAGGTCCTCAATCAAGATCAAAATCAA 61
QY 3419 GATCAGATCAAGATCAAGTTCCTCAATTTCACTGTGTTAGTCT 3464
Db 60 GATCAGATCAAGATCAAGTTCCTCAATTTCACTGTGTTAGTCT 15

RESULT 3
B10052/c
LOCUS
DEFINITION F23J11-Sp6 IGF Arabidopsis thaliana genomic clone F23J11, genomic survey sequence.
ACCESSION B10052
VERSION B10052.1 GI:2091184
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
/ eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 928)
AUTHORS Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and Ecker, J.
TITLE BAC End Sequences at ATGC
JOURNAL Unpublished
COMMENT Other_GSSs: F23J11-Sp6.1, F23J11-T7.1, F23J11-T7
Contact: Ecker J.
Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@atgenome.bio.upenn.edu
Seq primer: Sp6
Class: BAC ends
High quality sequence start: 105
High quality sequence stop: 133.
Location/Qualifiers
1..928
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="F23J11"
/sex="hermaphrodite"
/clone_lib="IGF"
/notes="Vector: BelobAC11; Site_1: EcoRI; Site_2: EcoRI; Produced by Thomas Altmann"

BASE COUNT 268 a 249 c 201 g 199 t 11 others
ORIGIN

Query Match 7.9%; Score 320.4; DB 28; Length 928;
Best Local Similarity 76.2%; Pred. No. 6e-37;
Matches 552; Conservative 0; Mismatches 151; Indels 21; Gaps 12;
QY 2739 TTTCTACAGTGAATGTGTGTTTTTATGTTTTTGTAGTTCTTTCACAGAGATGCTCCCAT 2798
Db 797 TTTCTACTTCACAAAGGAGTGGTGTGTTTAAATGTACGGTTCGGACGAGGCTCCCAT 738
QY 2799 CATATTAGTCTTTTCTTTTCTTTTGTGTGTTCTTCTCTGTAATGTTTCTTCTGATA 2858
Db 737 CAATTTTCCCTTCTCTCTCGGTCG---TGTAATGGGGAATTCGGTTCGGTA 684
QY 2859 AAGCTTTACTTCTTAAGTTTTTCCAGCGAGCGTGAATTTTACGTATCAACTCAATA 2918
Db 683 AAGCTTGACATTC-TCACTTTTTTCCAGCGAGCTGGA--TAATAGCGTATCATGCTCAGGA 627
QY 2919 TCCGACCTATAATTTAAAGAACACATTTAGCTAGATGTTTCACTTTTGAAAAATTTATTTCTAT 2978
Db 626 TCGACCTATAAA--TAACCGACACATAGCTAGAGTTTCACTCTTGAAGATTTATGCTAT 569
QY 2979 CGGAAGGGGGCAGAT-GAATTTCTGAGCATGTCTCTCGATCAATTTACCGTGAATACAAATG 3037
Db 568 CGGAAGGGGGTAAGATGATGATTTCTGAGCATGTCTCTCGATCAATTAACGTAGTGAATGG 509
QY 3038 TATTGCGTTGTTGT-TTCTAGATCTGGCTCACTGTCTACAGCTGGGATGGATCTAGA 3096
Db 508 TATCGCGTTGCTTGTNTTCTAGATCAGGCTCACTGTCTACGAGCTGGGATGGATCTAGA 449
QY 3097 TGGGTCACTAGATGGATTTCTGAGCTGGATTTACAAAGCTGGATTTAGCATGAC-TGAA 3155
Db 448 TGGTTCATCAAGATGGATTCGTGAGCTGGATCCACAGAGCTGGCTTCGCCCTGAATTTGAA 389
QY 3156 CTTCTGTTT-TACGGTCTGGTCTGGTCTGATCTCCGCGGTATCAGCTGTAGGATCTGA 3214
Db 388 CTTCTGTTCTACGGTCTGGTCTGGTCTGATCTCCGCGGTATCAGCTGTAGGATCTGA 329
QY 3215 TCGCAAAAGTTTGGACTATGATTTACTCTGATTTCTCTCAATATATTTATTTTTGACAATA 3274
Db 328 TGGCAAAAGCTCTGGACATGATTTACTCTGATTCGTGATATATTTGATCTTTTGGACAATA 269
QY 3275 GTGATTTCTGTTGTAGTCTTTTCTAGGACAGATTTAAGCTCCCGGAGTAGATGGGA 3334
Db 268 GTGACTCTGGGTGCG-GTTATATTATAGGACAGCATTTAAGCTCCCGGACTAGATGGCGA 210
QY 3335 GATGGTCAGTAAATTTCTTTGTTATGCCACACTACATGAGGTTTTTTCGGTCTGTGCGCA 3394
Db 209 TGT-----CAGTCAATTTGTTGTTATGCCACACTGACATGGGGTTTTT-GGTCTTGTGCA 155
QY 3395 GGTCCCAATCAAGATCAAAATCAAGATCAAGATCAAGATTTCTCCAGTTTCACTG 3454
Db 154 TGTCCCACTCGAGATCTAAGTCGAGATCAAGATCAAGATCGAATTCGAGCTCGTACCCG 95
QY 3455 TGGT 3458
Db 94 GGAT 91

RESULT 4
B11569
LOCUS
DEFINITION F3E19-Sp6 IGF Arabidopsis thaliana genomic clone F3E19, genomic survey sequence.
ACCESSION B11569
VERSION B11569.1 GI:2092690
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
/ eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 973)
AUTHORS Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and Ecker, J.

Qy 4015 GGTGTCCTTTATG 4027
Db 13 GGTGTCCTTTATG 1

RESULT 6
AV531252/c
LOCUS AV531252
DEFINITION Arabidopsis thaliana linear EST 01-SEP-2000
thaliana cDNA clone FB018C01F 3', mRNA sequence.

ACCESSION AV531252
VERSION AV531252
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 510)
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL DNA Res. 7, 175-180 (2000)
MEDLINE 20363093
PUBMED 10907847

COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp; URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1. 510
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="FB018C01F"
/tissue_type="flower buds"
/clone_lib="Arabidopsis thaliana flower buds Columbia"
/note="Vector: pBluescriptII SK-, Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 175 a 110 c 99 g 126 t

Query Match 6.7%; Score 272.4; DB 9; Length 510;
Best Local Similarity 99.3%; Pred. No. 6.9e-30;
Matches 284; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 3174 GTCTGCTGTGCTCCGCGGTATCAGCTGTAGGATCTGTCGCAAGTTTGGACTAT 3233
Db 510 GTCTGCTGTGCTCCGCGGTATCAGCTGTAGGATCTGTCGCAAGTTTGGACTAT 451

Qy 3234 GATTACTGTATCTCTCAATATTATTTTTCACCAATAGTGGANTCTGTGTGAGTT 3293
Db 450 GATTACTGTATCTCTCAATATTATTTTTCACCAATAGTGGANTCTGTGTGAGTT 391

Qy 3294 CTTTCTAGGACAGATTAACTCCGGGACTAGATCGGAGTGGTCAAGTAATTTCTT 3353
Db 390 CTTTCTAGGACAGATTAACTCCGGGACTAGATCGGAGTGGTCAAGTAATTTCTT 331

Qy 3354 TGTATGACACTTACATGGGTTTTCGGTCTTCTGCTGCGAGTCCCAATCAAGTCAA 3413
Db 330 TGTATGACACTTACATGGGTTTTCGGTCTTCTGCTGCGAGTCCCAATCAAGTCAA 272

Qy 3414 ATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAG 3459
Db 271 ATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAG 226

RESULT 7
BZ358568
LOCUS BZ358568

DEFINITION SALK_132986.49.85.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_132986.49.85.x, genomic
survey sequence.

ACCESSION BZ358568
VERSION BZ358568.1
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 276)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,
C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,I., Shinn,P.,
Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
JOURNAL Unpublished
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu

FEATURES
source
1. 276
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_132986.49.85.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/notes="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT 70 a 56 c 48 g 102 t

Query Match 6.8%; Score 265; DB 29; Length 276;
Best Local Similarity 99.6%; Pred. No. 1e-28;
Matches 276; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 658 CCAACGACGAGTGTGTCTCTCTCTCCAGCAATCAACGCTTCTCGAATCTTCT 717
Db 1 CCAACGACGAGTGTGTCTCTCTCTCCAGCAATCAACGCTTCTCGAATCTTCT 60

Qy 718 TCTTCTCTCTCTCGAAATTTTTCAGTAAATTTTTCAGTAAATTTTTCAGTAAATTTT 777
Db 61 TCTTCTCTCTCTCGAAATTTTTCAGTAAATTTTTCAGTAAATTTTTCAGTAAATTTT 120

Qy 778 AGGAACAAATTTTCTCTCTGAGGATCAGATGAGTGGGATCGAATCGTACGATCTAC 837
Db 121 AGGAACAAATTTTCTCTCTGAGGATCAGATGAGTGGGATCGAATCGTACGATCTAC 180

Qy 838 GTTGGGAATTTGCTCGAGATATTCGAAAGTGTGAGGTTGAAGATCTCTTCAAGGTT 897
Db 181 GTTGGGAATTTGCTCGAGATATTCGAAAGTGTGAGGTTGAAGATCTCTTCAAGGTT 240

Qy 898 TGAAGATTTCTCTCTCTCGATATAAATTTGAATT 934
Db 241 TGAAGATTTCTCTCTCTCGATATAAATTTGAATT 276

RESULT 8
B11684

LOCUS B11684 1044 bp DNA linear GSS 14-MAY-1997
DEFINITION F7G19-Sp6.2 IGF Arabidopsis thaliana genomic clone F7G19, genomic survey sequence.
ACCESSION B11684
VERSION B11684.1 GI:2092807
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (Chale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 1044)
REFERENCE Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and Ecker,J.
AUTHORS BAC End Sequences at ATGC
TITLE BAC End Sequences at ATGC
JOURNAL Unpublished
COMMENT Other GSSs: F7G19-T7.2, F7G19-T7.1, F7G19-T7.4, F7G19-Sp6.3, F7G19-T7.3, F7G19-T7, F7G19-Sp6, F7G19-Sp6.1
Contact: Ecker J.
Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@at.genome.bio.upenn.edu
Seg primer: Sp6
Class: BAC ends
High quality sequence start: 207
High quality sequence stop: 432.
Location/Qualifiers
1..1044
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="F7G19"
/sex="hermaphrodite"
/clone_lib="IGF"
/notes="Vector; BoloBACII; Site_1: EcoRI; Site_2: EcoRI; Produced by Thomas Altman"
BASE COUNT 297 a 228 c 107 g 341 t 71 others
ORIGIN
Query Match 6.5%; Score 263.4; DB 28; Length 1044;
Best Local Similarity 71.0%; Pred. No. 1e-28;
Matches 370; Conservative 0; Mismatches 145; Indels 6; Gaps 4;
Qy 3409 TCAAAATCAAGATCAAGATCAAGATCGAATTCACGTTTCACGTTGCGTAAGTCTAAAA 3468
Db 111 TCATATGATGCCNGGTACCAACTCAAAATCTCCATTTTCAGCT--GTAACCTCAAAA 167
Qy 3469 GCTGAACCTTCTTTAATTCACATCCAT-GTGTTTGTTTAAATACCTGCTCATTGGTT 3527
Db 168 NCTAAACCTTCTTTAAATTCACATCCATNGTGTGTGTTTAAATACCTGCTCATTGGTT 227
Qy 3528 GTTCTTCAATCAACACCAACTTAACGAAATCATGAGACAGACTATAAAATTCGAAGATC 3587
Db 228 GTTCTTCAATCAACACCAACTTAACGAAATCATGAGACAGACTATAAAATTCGAAGATC 287
Qy 3588 TGTAGAACGACTAGGTCTCACCAACCTCTGTGTGCACTATAAAATTCGCTCTCAAGTGT 3647
Db 288 T-FAGAACGACTANGTCTACCAACCTCTGTGTGCACTATAAAATTCGCTCTCAAGTGT 346
Qy 3648 TCAGCACATAATCTACTCTGTCATGTGTTATCATTTCTTCTCTTAAACGGTATTAC 3707
Db 347 TCAGCAACATAATCTACTCTGTCATGTGTTATCATTTCTTCTCTTAAACGGTATTAC 406
Qy 3708 ATATTATGTTTTCGAGGTGATATCTGGTTGAAATGAAGAACTGGCCACTGGCTGACC 3767
Db 407 ATATTATGTTTTCGAGGTGATACNNNNNNNANNTANNNNNNNNNNNNNNNNNNNC 466
Qy 3768 AATCGTCTCAAGCTTCTTCAGCGCTCCACTGCTGATAGAAATTTGATTCGATTTCCGATTT 3827

QY 3605 TCACCAACCTCTGTGTGCACTAAATCGCTTCCAGTGTTCAGCAACATAATCTAC 3664
Db 281 TCACCAACCTCTGTGTGCACTAAATCGCTTCCAGTGTTCAGCAACATAATCTAC 340
QY 3665 CTCGTGTCGTGTATCATCTTCTCTTAAAGGATTACATATATATTTTGCAGG 3724
Db 341 ATCTGTCATGTATCAATCTTCTCTTAAAGGATTACATATATATTTTGCAGG 400
QY 3725 TGATATCTCGTTGAAATGAAACCTGGCACCTGGCTGTACCCGAATCTCTCAAGCTTCT 3784
Db 401 TGATAAAGNN 459
QY 3785 CAGGCTCCACTCTCTATAGAAATTTGATCCGATTTGGGATTTATATATCTGTTCTTCTGT 3844
Db 460 --NGGCCNCCNNNNNCAATGNGACTCCGATCTGGGAAT-TCTCCCGTCTCTCTTT 516
QY 3845 ATGGGACGACCA 3856
Db 517 TGGGACACCA 528
RESULT 10
BZ592759
LOCUS 340 bp DNA linear GSS 07-JAN-2003
DEFINITION SALK_029105.23.05.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_029105.23.05.x, genomic
survey sequence.
ACCESSION BZ592759
VERSION BZ592759.1 GI:27533278
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 340)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.,
Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
JOURNAL Unpublished
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA.
Class: TDNA tagged.
Location/Qualifiers
1. 340
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_029105.23.05.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"
BASE COUNT 96 a 69 c 73 g 102 t
ORIGIN
Query Match 6.4%; Score 260.2; DB 29; Length 340;
Best Local Similarity 87.1%; Pred No. 4,7e-28;
Matches 297; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

QY 449 GATTGGAATTTGGTTCGAATGTGTAATAAATAAATAATTCCTCGGTTGAATGATAAG 508
Db 1 GATTGGAATTTGGTTCGAATGTGTAATAAATAAATAATTCCTCGGTTGAATGATAAG 60
QY 509 AAAAAACACATTCGATCCGGTTAGAACATATATTAACAGGCCCATTTAAACATATGGGC 568
Db 61 AAAAAACACATTCGATCCGGTTAGAACATATATTAACAGGCCCATTTAAACATATGGGC 119
QY 569 CGATCTTTGATCAACTGGGCTTATTCATCGTTTGATACATCGCGGCCACAGGATTTAAATCC 628
Db 120 CGATCTTTGATCAACTGGGCTTATTCATCGTTTGATACATCGCGGCCACAGGATTTAAATCC 179
QY 629 AGTTCCGTTTATAAAGGATAGTATTTCCAAACGAAACGGTGGTGTCTCTCTTTCCAGA 688
Db 180 GGTGCCGTTTATAAAGGATAGTATTTCCAAACGAAACGGTGGTGTCTCTCTTTCCAGA 239
QY 689 ACATCTTAACGCTTCTCGAACATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 748
Db 240 ACATCTTAACGCTTCTCGAACATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 299
QY 749 AATCAATTTCTTCT 789
Db 300 AATCAATTTCTTCT 340
RESULT 11
B12780
LOCUS 699 bp DNA linear GSS 14-MAY-1997
DEFINITION F7G19-Sp6.1 IGF Arabidopsis thaliana genomic clone F7G19, genomic
survey sequence.
ACCESSION B12780
VERSION B12780.1 GI:2093901
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 699)
AUTHORS Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and
Ecker,J.
TITLE BAC End Sequences at ATGC
JOURNAL Unpublished
COMMENT Other GSSs: F7G19-T7.2, F7G19-T7.1, F7G19-T7.4, F7G19-Sp6.3,
F7G19-T7.3, F7G19-Sp6.2, F7G19-T7, F7G19-Sp6
Contact: Ecker J.
Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA
19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@genome.bio.upenn.edu
Seq primer: Sp6
Class: BAC ends
High quality sequence start: 113
High quality sequence stop: 474.
FEATURES
source
1. 699
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="F7G19"
/sex="hermaphrodite"
/clone_lib="IGF"
/note="Vector: BelobACII; Site_1: EcoRI; Site_2: EcoRI;
Produced by Thomas Altmann"
BASE COUNT 103 a 235 c 77 g 274 t 10 others
ORIGIN
Query Match 6.1%; Score 247.8; DB 28; Length 699;

Best Local Similarity 94.5%; Pred. No. 2.2e-26;	
Matches 290; Conservative 0; Mismatches 12; Indels 5; Gaps 3;	
Qy	3426 ATCAAGATCGAATTTCTTCAGTTTCACCTGTGGTAAGTCTAAAGCTTAAAGCTGAACCTTCTTTTAAT 3485
Db	90 ACCGAGCTCGAGTCTCTCCCTTTTCACTG---TAAATCTAAAGCTGAACCTTCTTTTAAT 146
Qy	3486 TCACAAATCCAT-GTGTTTGGTTTAAATACCTGCTCACTTTGGTTTCTTCAATCAACACC 3544
Db	147 TCACAAATCCATGGTGTGTGTGTTTAAATACCTGCTCACTTTGGTTTCTTCAATCAACACC 206
Qy	3545 AACTTAACGAAATCATGAGCAGACTATAAAATTTGAAGAGTCTGTAGACGACTAGGTC 3604
Db	207 AACTTAACGAAATCATGAGCAGACTATAAAATTTGAATAGTCT-TAGAACGACTAGGTC 265
Qy	3605 TCACCAACCTCTGTGTGCACATAAAATTCGCCTCTCCAAAGTGTTCAGCAACATAATCTAC 3664
Db	266 TCACCAACCTCTGTGTGCATAAAATTCGCCTCTCCAAAGTGTTCAGCAACATAATCTAC 325
Qy	3665 CTCGTGATGTTTATCATTTCTTCTTCTTAAACGGTATTACATATATTGTTTTGCAGG 3724
Db	326 CTCGTGATGTTTATCATTTCTTCTTCTTAAACGGTATTACATATATTGTTTTGCAGG 385
Qy	3725 TGATATC 3731
Db	386 TGATAAC 392
RESULT 12	
BZ471603/c	
LOCUS	BZ471603
DEFINITION	BONJMW74R BO 1.6 2 KB tot Brassica oleracea genomic clone BONJW74,
ACCSSION	genomic survey sequence.
VERSION	BZ471603
KEYWORDS	BZ471603.1 GI:26769878
SOURCE	GSS.
ORGANISM	Brassica oleracea
REFERENCE	Brassica oleracea
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
JOURNAL	; eurosids II; Brassicales; Brassicaceae; Brassica.
COMMENT	1 (bases 1 to 791)
	Whole genome shotgun sequencing of Brassica oleracea
	Unpublished
	Other GSSs: BONJW74TF
	Contact: Chris Town
	TIGR
	9712 Medical Center Drive, Rockville, MD 20850, USA.
	Tel: 301-838-3523
	Fax: 301-838-0208
	Email: cdtown@tigr.org
	DNA is from a doubled haploid provided by Tom Osborn.
	Seq primer: TR
	Class: Sheared ends.
FEATURES	Location/Qualifiers
source	1..791
	/organism="Brassica oleracea"
	/mol_type="genomic DNA"
	/strain="Tol000DH3"
	/db_xref="taxon:3712"
	/clone="BONJW74"
	/clone_lib="BO 1.6 2 KB tot"
	/note="vector: phosI; Site 1: BstXI; 1.6-2 kb sheared
	total DNA inserted into phosI using BstXI linkers"
BASE COUNT	295 a 185 c 119 g 192 t
ORIGIN	
Query Match	6.1%; Score 246.4; DB 29; Length 791;
Best Local Similarity	64.3%; Pred. No. 3.4e-22;
Matches 491; Conservative	0; Mismatches 221; Indels 52; Gaps 6;
Qy	1523 CGGATTACCGCCCTTCTGCTTGGCGAGGACCTTAAGGTAAGGCACACTATAGTCTTT 1582

AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Rosso, M., Strishov, N., Li, Y., Reiss, B., Dekker, K. and Weishaar, B.	A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)		3 (bases 1 to 298)	Strishov, N., Rosso, M., Li, Y. and Weishaar, B.			
	for flanking sequence tag based reverse genetics		Unpublished		Direct Submission		
			Submitted (21-OCT-2002)	Weishaar B., Max-Planck-Institut fuer			
			Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany				
			indicates an insertion close to the left border of the T-DNA. It				
			sequences are generated at the MPI for Plant Breeding Research in				
			the context of the GABI-Kat project. GABI-Kat is part of the German				
			Plant Genomics program designated 'GABI'. Information on line				
			availability can be found at:				
			http://www.mpiz-koeln.mpg.de/GABI-Kat/ .				

```

FEATURES
source
Location/Qualifiers
1..238
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="GK-325E11-015977"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/notes="PCR was performed on DNA from Arabidopsis thaliana plants (r1) which were transformed with the T-DNA from vector pAC151. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thaliana nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"

```

BASE COUNT	74 a	44 c	70 g	110 t
ORIGIN				
Query Match		6.0%;	Score 243.4;	DB 29;
Best Local Similarity		99.8%;	Pred. No. 1.4e-25;	Length 298;
Matches 244;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;

Qy	1026	CAATTGGGACATTGATTTGAAGATTCCACCAGAGACTCTCGTTATGCTTTGTGCGAGG	1085
D _b	35	CAATTGAGGACATTGATTTGAAGATTCCACCAGAGACTCTCGTTATGCTTTGTGCGAGG	94

QY 1086 TATATTGATCAAGTACAAAATTGTTTTTCTTCTCTTGTAAATAGTATAGGCTAATGAC 1145

DB		95	TATATTGAACCAAGTACAAATTTGTTTTCTTCCTTGTAATAGTATAGGGCTAATGAC	154
QV		1146	TAAGATAGTTTTGTATTGGTGGCAGTTTTGAAGATCCCTCGTGGATGCAGACGATGCCAATTTA	1205

Db 155 TAAGATAGTTTGTATTGGTGGCAGTTGAAGATCCTCGTATGCAGACGATGCAATTTA 214

1206	1205	1204	1203	1202	1201	1200	1199	1198	1197	1196	1195	1194	1193	1192	1191	1190	1189	1188	1187	1186	1185	1184	1183	1182	1181	1180	1179	1178	1177	1176	1175	1174	1173	1172	1171	1170	1169	1168	1167	1166	1165	1164	1163	1162	1161	1160	1159	1158	1157	1156	1155	1154	1153	1152	1151	1150	1149	1148	1147	1146	1145	1144	1143	1142	1141	1140	1139	1138	1137	1136	1135	1134	1133	1132	1131	1130	1129	1128	1127	1126	1125	1124	1123	1122	1121	1120	1119	1118	1117	1116	1115	1114	1113	1112	1111	1110	1109	1108	1107	1106	1105	1104	1103	1102	1101	1100	1099	1098	1097	1096	1095	1094	1093	1092	1091	1090	1089	1088	1087	1086	1085	1084	1083	1082	1081	1080	1079	1078	1077	1076	1075	1074	1073	1072	1071	1070	1069	1068	1067	1066	1065	1064	1063	1062	1061	1060	1059	1058	1057	1056	1055	1054	1053	1052	1051	1050	1049	1048	1047	1046	1045	1044	1043	1042	1041	1040	1039	1038	1037	1036	1035	1034	1033	1032	1031	1030	1029	1028	1027	1026	1025	1024	1023	1022	1021	1020	1019	1018	1017	1016	1015	1014	1013	1012	1011	1010	1009	1008	1007	1006	1005	1004	1003	1002	1001	1000	999	998	997	996	995	994	993	992	991	990	989	988	987	986	985	984	983	982	981	980	979	978	977	976	975	974	973	972	971	970	969	968	967	966	965	964	963	962	961	960	959	958	957	956	955	954	953	952	951	950	949	948	947	946	945	944	943	942	941	940	939	938	937	936	935	934	933	932	931	930	929	928	927	926	925	924	923	922	921	920	919	918	917	916	915	914	913	912	911	910	909	908	907	906	905	904	903	902	901	900	899	898	897	896	895	894	893	892	891	890	889	888	887	886	885	884	883	882	881	880	879	878	877	876	875	874	873	872	871	870	869	868	867	866	865	864	863	862	861	860	859	858	857	856	855	854	853	852	851	850	849	848	847	846	845	844	843	842	841	840	839	838	837	836	835	834	833	832	831	830	829	828	827	826	825	824	823	822	821	820	819	818	817	816	815	814	813	812	811	810	809	808	807	806	805	804	803	802	801	800	799	798	797	796	795	794	793	792	791	790	789	788	787	786	785	784	783	782	781	780	779	778	777	776
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Qy 1266 AGCTA 1270
 .|.|.|.|

Db 275 AGCTA 279

RESULT 14
AV532256/C
LOCUS

DEFINITION AV532256 *Arabidopsis thaliana* flower buds Columbia
thaliana CDNA clone FB039a11F 3', mRNA sequence.

VERSION AV532256.1 GI:8692539
KEYWORDS EST.

ORGANISM

```

; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
  1 (bases 1 to 300)
AUTHORS
  Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE
  A large scale analysis of cDNA in Arabidopsis thaliana: Generation
  of 13,028 non-redundant expressed sequence tags from normalized and
  size-selected cDNA libraries
JOURNAL
  DNA Res. 7, 175-180 (2000)
MEDLINE
  20363093
PUBMED
  10907847
COMMENT
  Contact: Erika Asamizu
  The First Laboratory for Plant Gene Research
  Kazusa DNA Research Institute
  Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
  Email: asamizukazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
  Location/Qualifiers
FEATURES
  source
    1..300

```

```

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="FB039a11F"
/tissue_type="flower buds"
/clone_lib="Arabidopsis thaliana flower buds Columbia"
/notes="Vector: pBluescript1 SK-, Site_1: EcoRI; Site_2:
XhoI"
95 a 60 c 62 g 83 t
BASE COUNT

```

ORIGIN	Query Match	5.8%	Score 235.6	DB 9	Length 300
--------	-------------	------	-------------	------	------------

Matches	238	Conservative	0	Mismatches	4	Indels	0	Gaps	0
Best Local Similarity	98.3%	Pred. NO.	1.8e-24						

Qy	3715	GTATTTGCAGTGATATCTCGTTTGAAAAATGAAAACTGGCCACTGCGCTGTACCCCGAATCGTC	3774
DP	243	GTATTTCCACCTGTTGATATCTCGTTTGAAAAATGAAAACTGGCCACTGCGCTGTACCCCGAATCGTC	184

QY 3775 TCAAGCTTCTCAGGCTCCACTGCTAATAGAAATTGATTCCGATTGGGATTATTATACTG 3834

Db	183	TCAAGCTTCTCAGGCTCCACTGCTAATAGAAATTGATTCGGATTATATTA	CTG 124
Qv	3835	GCCTTCTTCTCAGGCTCCACTGCTAATAGAAATTGATTCGGATTATATTA	CTG 3894

Db 123 GTCTTCCTGTATGGGACGACCAATATGTCTTTCTAGTTTAGTTGTGAACCTGGAATTGG 64

Qy	3895	TCGTATTGTGTCATTAAAAAGCCGGAAAC	TCGTCTCGGCTGCATAATAAAGTTCATC	3954
+	3	TCGTATTGTGTCATTAAAAAGCCGGAAAC	TCGTCTCGGCTGCATAATAAAGTTCATC	4

Qy 3955 AG 3956
11.

3 AG 2
Db

RESULT 15
AI998038/C

DEFINITION
701672742 A. thaliana, Columbia Col-0, rosette-1 Arabidopsis
thaliana cdna clone 701672742, mRNA sequence.

ACCESSION AI398038
 VERSION AI998038.1 GI:5845023
 KEYWORDS EST.

ORGANISM	SOURCE
<i>Arabisopsis thaliana</i>	Arabidopsis thaliana (Crucif. Cress)
<i>Arabisopsis thaliana</i>	Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Embryophyta; Tracheophyta;	

REFERENCE
1 (bases 1 to 439)
; eucosids II; Brassicales; Brassicaceae; Arabidopsis.
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosales

AUTHORS

Chen, J., Momiyama, M., Chan, S., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Broska, P., Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobriga, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and Hanson, D.

TITLE Arabidopsis thaliana Gene Expression MicroArray
JOURNAL Unpublished
COMMENT Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte
Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: service@genomesystems.com.

FEATURES Location/Qualifiers
source 1..439
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /cultivar="Columbia Col-0"
 /db_xref="taxon:3702"
 /clone="701672742"
 /tissue_type="rosette"
 /dev_stages="4 - 7 weeks"
 /clone_lib="A. thaliana, Columbia Col-0, rosette-1"
 /notes="Vector: pSPORT; Site 1: NotI; Site 2: SalI; cDNA library was derived from untreated rosette tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was bluntended, ligated to SalI adaptors, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."
BASE COUNT 154 a 91 c 81 g 113 t
ORIGIN

Query Match 5.6%; Score 227.4; DB 9; Length 439;
Best Local Similarity 99.2%; Pred. No. 2.4e-23;
Matches 239; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 3219 AAGTTTGGACTATGATTCGATTCCTCAATATATTTATCTTTTGACAAATAGTG 3278
Db |||||||
QY 439 AAGTTTGGACTATGATTCGATTCCTCAATATATTTATCTTTTGACAAATAGTG 380
Db |||||||
QY 3279 ATTCTGTCTTGAAGTTCTTTCTAGGACAGCATTTAAGCTCCGGGACTAGATGGAGATG 3338
Db |||||||
QY 379 ATTCTGTCTTGAAGTTCTTTCTAGGACAGCATTTAAGCTCCGGGACTAGATGGAGATG 320
Db |||||||
QY 3339 GTCAGTAAATTTCTTTGTTATGCCACACTTACATGGGGTTTTTCGGTCTTGTCCAGTTC 3398
Db |||||||
QY 319 GTCAGTAAATTTCTTTGTTATGCCACACTTACATGGGGTTTTT-GGTCTTGTCCAGTTC 261
Db |||||||
QY 3399 CCAATCAAGATCAAAATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGAT 3458
Db |||||||
QY 260 CCAATCAAGATCAAAATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGAT 201
Db |||||||
QY 3459 A 3459
Db 200 A 200

Search completed: January 29, 2004, 02:49:17
Job time : 7856 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 28, 2004, 16:39:25 ; Search time 216 Seconds
(without alignments)
8263.673 Million cell updates/sec

Title: US-10-014-927-18
Perfect score: 4044
Sequence: 1 agacaaagatgcttacttct.....atgattatccttaactaac 4044

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	78.6	1.9	5285	US-09-402-328-1	Sequence 1, Appli
C 2	75.8	1.9	7218	US-08-232-463-14	Sequence 14, Appl
C 3	54.4	1.3	7218	US-08-232-463-14	Sequence 14, Appl
C 4	48.8	1.2	193	US-09-702-705-1719	Sequence 1719, Ap
C 5	48.8	1.2	193	US-09-736-457-1719	Sequence 1719, Ap
C 6	48.2	1.2	284	US-09-313-294A-5971	Sequence 5971, Ap
C 7	47.8	1.2	19124	US-08-487-826B-13	Sequence 13, Appl
C 8	45.8	1.1	2447	US-09-014-969-14	Sequence 14, Appl
C 9	41.8	1.0	1117	US-09-247-373B-33	Sequence 33, Appl
C 10	41.4	1.0	1586	US-09-152-060-12	Sequence 12, Appl
C 11	41.4	1.0	2773	US-09-996-243-178	Sequence 178, App
C 12	40.8	1.0	5511	US-08-928-361B-2	Sequence 2, Appli
C 13	40.8	1.0	5511	US-09-588-995A-2	Sequence 2, Appli
C 14	40.8	1.0	7334	US-08-928-361B-1	Sequence 1, Appli
C 15	40.8	1.0	7334	US-09-588-995A-1	Sequence 1, Appli
C 16	40.6	1.0	658	US-08-998-416-595	Sequence 595, App
C 17	40.6	1.0	5163	US-08-700-651-1	Sequence 1, Appli
C 18	40.6	1.0	5163	US-08-928-361B-4	Sequence 4, Appli
C 19	40.6	1.0	5163	US-09-588-995A-4	Sequence 4, Appli
C 20	40.6	1.0	5318	US-08-700-651-2	Sequence 2, Appli
C 21	40.6	1.0	5318	US-08-928-361B-3	Sequence 3, Appli
C 22	40.6	1.0	5318	US-09-588-995A-3	Sequence 3, Appli
C 23	40.4	1.0	1798	US-09-797-906-1	Sequence 1, Appli
C 24	39.6	1.0	665	US-08-998-416-937	Sequence 937, App
C 25	39.6	1.0	1664976	US-08-916-421B-1	Sequence 1, Appli
C 26	39.4	1.0	240	US-08-628-417-6	Sequence 6, Appli
C 27	39.2	1.0	6755	US-08-931-999-4	Sequence 4, Appli

28	39	1.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
C 29	38.6	1.0	662	3	US-08-998-416-185	Sequence 185, App
C 30	38.6	1.0	701	3	US-08-998-416-701	Sequence 701, App
C 31	38.6	1.0	711	3	US-08-998-416-786	Sequence 786, App
C 32	38.6	1.0	724	3	US-08-998-416-683	Sequence 683, App
C 33	38.6	1.0	722	3	US-08-998-416-1036	Sequence 1036, App
C 34	38.6	1.0	767	3	US-08-998-416-472	Sequence 472, App
C 35	38.6	1.0	782	3	US-08-998-416-224	Sequence 224, App
C 36	38.6	1.0	827	3	US-08-998-416-535	Sequence 535, App
C 37	38.6	1.0	828	3	US-08-998-416-538	Sequence 538, App
C 38	38.6	1.0	834	3	US-08-998-416-305	Sequence 305, App
C 39	38.2	0.9	4395	2	US-08-849-480A-3	Sequence 3, Appli
C 40	38	0.9	494	2	US-08-332-766A-22	Sequence 22, Appli
C 41	38	0.9	1696	4	US-09-835-811-1	Sequence 1, Appli
C 42	38	0.9	5852	1	US-07-867-106-2	Sequence 2, Appli
C 43	38	0.9	6243	2	US-09-056-075-1	Sequence 1, Appli
C 44	37.8	0.9	10640	4	US-09-417-485D-5	Sequence 5, Appli
C 45	37.4	0.9	289	3	US-09-007-005-17	Sequence 17, Appli

ALIGNMENTS

RESULT 1
US-09-402-328-1/c
; Sequence 1, Application US/09402328
; Patent No. 6365728
; GENERAL INFORMATION:
; APPLICANT: Purdue Research Foundation,
; Hodges, Thomas K.
; Lysnik, Leszek A.
; TITLE OF INVENTION: Regulatory Element For Expressing Genes
; IN Plants
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Barnes & Thornburg
; STREET: 11 S. Meridian
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/402,328
; FILING DATE: 05-No. 6365728-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Breen, John P.
; REGISTRATION NUMBER: 38,833
; REFERENCE/DOCKET NUMBER: 3220-29933
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 231-7745
; TELEFAX: (317) 231-7433
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5285 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Arabidopsis thaliana
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-402-328-1

Query Match 1.9%; Score 78.6; DB 4; Length 5285;
Best Local Similarity 51.1%; Pred. No. 5.8e-11;

APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-F1s
US-08-232-463-14

Query Match 1.3%; Score 54.4; DB 1; Length 7218;
Best Local Similarity 3.3%; Pred. No. 0.00024;
Matches 13; Conservative 224; Mismatches 155; Indels 0; Gaps 0;
QY 81 CTGCAAGCAACAGTGGAGAGATTGATTAACGGAATATCGGAACATCACTCAACAAC 140
1435 CRR 1376
Db
QY 141 CAAAAATTTGGACATCATATCGCAACAAATTCATAGGAAAAATCTGAAATTCACAAAC 200
1375 RRR 1316
Db
QY 201 AGAAAAACCAACGGAACAGACAGCACTACGAGTACGAGACACCGTGACGGTGC 260
1315 RRR 1256
Db
QY 261 ACGCGGAAAAATGCTTGGACGGAGTACTAATCGGCAATTGAGATTGAGAGTGGT 320
1255 RRR 1196
Db
QY 321 AGTAGAGGAACGAGAGAAATGTTCTCTCAAAAAATCCCAAGTGTTCGATCTAGTG 380
1195 RRR 1136
Db
QY 381 TCTCTTTTGTCAAAACAGCAGTGTTCAGGAAACCTAGGAAATGAATGACCCGATGT 440
1135 RRR 1076
QY 441 CCGAATCCGATCGAAATGGTTCCAAATGTA 472

Db 1075.RRRRRRRRATCGAAGCTCCCTCGACCTGCA 1044
RESULT 4
US-09-702-705-1719
Sequence 1719, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1719
LENGTH: 193
TYPE: DNA
ORGANISM: Homo sapiens
US-09-702-705-1719
Query Match 1.2%; Score 48.8; DB 4; Length 193;
Best Local Similarity 77.6%; Pred. No. 0.0011;
Matches 59; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 1172 TTGAAGATCCTCTGTGTCGAGACGATGCAATTTATGACGTGATGATTGATTTGATG 1231
Db 1 TCGAGGACCCCGAGATGTCGAGAGATGCTATTATGGAAGAAATGTTATGATTATGCC 60
QY 1232 GGTGTCGACTTCGGGT 1247
Db 61 AGTGTGCGTTCGTGT 76

RESULT 5
US-09-736-457-1719
Sequence 1719, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1719
LENGTH: 193
TYPE: DNA
ORGANISM: Homo sapiens
US-09-736-457-1719
Query Match 1.2%; Score 48.8; DB 4; Length 193;

SULT 7
-08-487-826B-13/c
Sequence 13, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellem, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 520 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:

RESULT 8
US-09-014-969-14/c
; Sequence 14, Application US/09014969
; Patent No. 5963397
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R.
; APPLICANT: Ragie, Lisa A.
; APPLICANT: Merberg, David

```

1  RESULT 10
2  US-09-152-060-12/c
3  ; Sequence 12, Application US/09152060
4  ; Patent No. 6448230
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Rosen et al.
7  ; TITLE OF INVENTION: 28 Human Secreted Proteins
8  ; FILE REFERENCE: P2003P1.US
9  ; CURRENT APPLICATION NUMBER: US/09/152,060
10 ; CURRENT FILING DATE: 1998-09-11
11 ; EARLIER APPLICATION NUMBER: PCT/US98/04858
12 ; EARLIER FILING DATE: 1998-03-12
13 ; EARLIER APPLICATION NUMBER: 60/040,762
14 ; EARLIER FILING DATE: 1997-03-14
15 ; EARLIER APPLICATION NUMBER: 60/040,710
16 ; EARLIER FILING DATE: 1997-03-14
17 ; EARLIER APPLICATION NUMBER: 60/050,934
18 ; EARLIER FILING DATE: 1997-05-30
19 ; EARLIER APPLICATION NUMBER: 60/048,100
20 ; EARLIER FILING DATE: 1997-05-30
21 ; EARLIER APPLICATION NUMBER: 60/048,357
22 ; EARLIER FILING DATE: 1997-05-30
23 ; EARLIER APPLICATION NUMBER: 60/048,189
24 ; EARLIER FILING DATE: 1997-05-30
25 ; EARLIER APPLICATION NUMBER: 60/057,765
26 ; EARLIER FILING DATE: 1997-09-05
27 ; EARLIER APPLICATION NUMBER: 60/048,970
28 ; EARLIER FILING DATE: 1997-06-06
29 ; EARLIER APPLICATION NUMBER: 60/066,368
30 ; EARLIER FILING DATE: 1997-12-19
31 ; NUMBER OF SEQ ID NOS: 118
32 ; SOFTWARE: PatentIn Ver. 2.0
33 ; SEQ ID NO 12
34 ; LENGTH: 1586
35 ; TYPE: DNA
36 ; ORGANISM: Homo sapiens

```

US-09-152-060-12

Query Match 1.0%; Score 41.4; DB 4; Length 1586;
Best Local Similarity 58.5%; Pred. No. 0.35;
Matches 72; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy	2286	TCTTTTTCATTAACGACTTCTAATGCAATGCAGTGTCACCTGTAGATCCATTTC	2345
Dd	1577	TTTTTTTTCATTAACGACTTCTAATGCAATGCAGTGTCACCTGTAGATCCATTTC	1518
Qy	2346	ATTTATGAATTAACGACTTCTAATGCAATGCAGTGTCACCTGTAGATCCATTTC	2405
Dd	1517	TTAAATATGGTATCTGATAACCAATAGTCCCTCCCATCATTTCTCTCTCGCATCCTGGTTC	1458
Qy	2406	CCC 2408	
Dd	1457	CAC 1455	

RESULT 11
US-09-996-243-178/c
Sequence 178, Application US/09996243
Patent No. 647825
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gottitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C13
CURRENT APPLICATION NUMBER: US/09/996,243
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087600
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801

```

; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 1.0%; Score 41.4; DB 4; Length 2773;
Best Local Similarity 54.2%; Pred. No. 0.48;
Matches 84; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 2728 TTCCTTTTCATCTTACACGGAATGGTTGTTTTTATGTTTTGAGTTCTTGACAGA 2787
Db 2770 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2711

QY 2788 GATGCTCCATCATATTAGTCCCTTTTCCTTTCTCTTTGTGTCGTTCTCTTCGATGTT 2847
Db 2710 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2851

QY 2848 TCCTTCTGATAAAGCTTTTACTTCTTAACTTTTTC 2882
Db 2650 TCATCCTCTTAACATTTTAATTCAGCAATTCCTAC 2616

RESULT 12
US-08-928-361B-2/c
; Sequence 2, Application US/08928361B
; Patent No. 6071519
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,062
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: VERNY, HANA
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480.76-1 (HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5511 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-928-361B-2

Query Match 1.0%; Score 40.8; DB 3; Length 5511;
Best Local Similarity 52.3%; Pred. No. 1;
Matches 90; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 2190 AGAGTGTGTCGAGCCCGAGATGTTCTAAAGCTATAGAACGAGTCGAGCCGT 2249
Db 1038 ACTAGTTGTTGTCGTAGTAGTAGTGTAGTGTGTTGTTGTTGTCGT 979

```

QY 2250 GGTCCAAGCTAGCTATAGTAGCAAGCAGGAGGCTCTTTTTCAT 2309
Db 978 TGTGTTGTTGTCGCTAGTAGTAGTAGTAGTTGTTGTTGTTGTTGTT 919
QY 2310 AACCTAAGACATATAGGAGTTTATTTGTAAGTATTTATGAATTAAT 2361
Db 918 CACTCCACACATTTAGTGTATGAATGTATACCTTACTGCCCATATCAAT 867

RESULT 13

US-09-588-995A-2/c
; Sequence 1, Application US/09588995A
; Patent No. 6514697
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: BARNES, DEBRA A.
; APPLICANT: NELSON, RICHARD C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND
; TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: INFECTIONS
; FILE REFERENCE: 480.19-5
; CURRENT APPLICATION NUMBER: US/09/588,995A
; PRIOR FILING DATE: 2000-06-06
; PRIOR FILING DATE: 1997-03-27
; PRIOR FILING DATE: 1997-03-27
; PRIOR FILING DATE: 1997-09-12
; PRIOR FILING DATE: 1996-08-14
; PRIOR FILING DATE: 1996-08-14
; PRIOR FILING DATE: 1995-04-03
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 5511
; TYPE: DNA
; ORGANISM: Cryptosporidium parvum
US-09-588-995A-2

Query Match 1.0%; Score 40.8; DB 4; Length 5511;
Best Local Similarity 52.3%; Pred. No. 1;
Matches 90; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
QY 2190 AGGAGTGTGAGTCGAGCCAGATCTCTAAAGCTATAGAGCAGGAGTGGAGCCGT 2249
Db 1038 AGTAGTTGTTGTCGCTAGTAGTAGTAGTTGTTGTTGTTGTTGTTGTT 979
QY 2250 GGTCCAAGCTAGCTATAGTAGCAAGCAGGAGGCTCTTTTTCAT 2309
Db 978 TGTGTTGTTGTCGCTAGTAGTAGTAGTAGTTGTTGTTGTTGTTGTTGTT 919
QY 2310 AACCTAAGACATATAGGAGTTTATTTGTAAGTATTTATGAATTAAT 2361
Db 918 CACTCCACACATTTAGTGTATGAATGTATACCTTACTGCCCATATCAAT 867

RESULT 14

US-08-928-361B-1/c
; Sequence 1, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto

STATE: CA
COUNTRY: USA
ZIP: 94306-1840
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,361B
FILING DATE: 12-SEP-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,062
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Verny, Hana
REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: 480,76-1 (HV)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-1677
TELEFAX: 650-324-1678
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 7334 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-928-361B-1

Query Match 1.0%; Score 40.8; DB 3; Length 7334;
Best Local Similarity 52.3%; Pred. No. 1.2;
Matches 90; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
QY 2190 AGGAGTGTGAGTCGAGCCAGATCTCTAAAGCTATAGAGCAGGAGTGGAGCCGT 2249
Db 2706 AGTAGTTGTTGTCGCTAGTAGTAGTAGTTGTTGTTGTTGTTGTTGTT 2647
QY 2250 GGTCCAAGCTAGCTATAGTAGCAAGCAGGAGGCTCTTTTTCAT 2309
Db 2646 TGTGTTGTTGTCGCTAGTAGTAGTAGTAGTTGTTGTTGTTGTTGTTGTT 2587
QY 2310 AACCTAAGACATATAGGAGTTTATTTGTAAGTATTTATGAATTAAT 2361
Db 2586 CACTCCACACATTTAGTGTATGAATGTATACCTTACTGCCCATATCAAT 2535

RESULT 15

US-09-588-995A-1/c
; Sequence 1, Application US/09588995A
; Patent No. 6514697
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: BARNES, DEBRA A.
; APPLICANT: NELSON, RICHARD C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND
; TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: INFECTIONS
; FILE REFERENCE: 480.19-5
; CURRENT APPLICATION NUMBER: US/09/588,995A
; CURRENT FILING DATE: 2000-06-06
; PRIOR FILING DATE: 1997-03-27
; PRIOR FILING DATE: 1997-03-27
; PRIOR FILING DATE: 1997-09-12
; PRIOR FILING DATE: 1997-09-12
; PRIOR FILING DATE: 1996-08-14
; PRIOR FILING DATE: 1996-08-14
; PRIOR FILING DATE: 1995-04-03
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: Patent in Ver. 2.1

```
; SEQ ID NO 1
; LENGTH: 7334
; TYPE: DNA
; ORGANISM: Cryptosporidium parvum
US-09-588-995A-1

Query Match      1.0%; Score 40.8; DB 4; Length 7334;
Best Local Similarity 52.3%; Pred. No. 1.2;
Matches 90; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY      2190 AGGAGTGTGAGTCGAGCCCGAGATGATTCCTAAAGCTATAGAACGACGAGAGTCGGAGCCCT 2249
Db      |||||
Db      2706 AGTAGTTGTTGTGTCCTAGTAGTAGTGGTAGTAGTTGTAGTAGTTGTGTTGTCGT 2647
QY      2250 GGTCCAAGCTGTAGCTATAGTCGACAGCAGGAGCTTTTTTTTTTTTTTTTTCAT 2309
Db      |||||
Db      2646 TGTGTGTTGTTGTGTCCTAGTAGTAGTAGTTGTGTTGTTGTTGTTGTTGTTGTT 2587
QY      2310 AAACCTAAGACATATAAGGATTTTATTGTAACCTTATTATGAAATTAAC 2361
Db      |||||
Db      2586 CACTCCAACACATTTAGTGTATGGAATGTATACCTTACTGCCCATATCAAT 2535

Search completed: January 29, 2004, 00:38:13
Job time : 219 secs
```



Creation date: 04-12-2004
Indexing Officer: SMOHAMMED - SUAD MOHAMMED
Team: 1600PrintWorkingFolder
Dossier: 10133973

Legal Date: 04-09-2004

No.	Doccode	Number of pages
1	CTRS	7

Total number of pages: 7

Remarks:

Order of re-scan issued on